4-2014

Where do I know that? A distributed multimodal model of semantic knowledge

Kevin M. Stubbs
Western University, kmstubbs@gmail.com

Where Do I Know That?

A Distributed Multimodal Model of Semantic Knowledge

Kevin. M. Stubbs.

Honors Psychology Thesis
Department of Psychology
University of Western Ontario
London, Ontario, CANADA
April, 2014

Thesis Advisors: Ken McRae, Ph.D., and Chris McNorgan, Ph.D.

Abstract

As computers have grown more and more powerful, computational modeling has become an increasingly valuable tool for evaluating real world findings. Likewise, brain imaging has become increasingly powerful as is evidenced by recent fMRI findings which support the exciting possibility that semantic memory is segregated by modality in the brain (Goldberg et al., 2006b). The present study utilizes connectionist modeling to put the distributed multi-modal framework of semantic memory to the test, and represents the next step forward in the line of sensory-functional models. This model, based around the McRae et al. (2005) feature production norms, includes individual implementations of each modality: visual colour, visual motion, visual form and surface, olfactory-gustatory, encyclopedic, tactile, auditory, and functional. A cross-modal convergence zone (Hub), visual decoding region, and abstracted implementations of wordforms and images are also included to ultimately simulate picture naming. Focal lesions are simulated in each semantic modality, successfully recreating various category-specific deficits, many of which have been reported in patient case studies. Categories are expanded from a living-nonliving dichotomy to include animals, artifacts (tools, utensils, containers, and clothing), fruits and vegetables, and musical instruments. Hub damage recreates semantic dementia with an additional slight but significant impairment of animals over artifacts.

Acknowledgments

I would like to thank Dr. Ken McRae for being such a wonderful professor and advisor. If it were not for Ken, I may never have ever discovered my keen interest in cognitive science and computational methods. I would also like to extend my most sincere thanks to Dr. Chris McNorgan for all of his help in brainstorming and troubleshooting during this project. Without the guidance of Ken and Chris, this project would simply not have been possible. I am looking forward to continuing work on this project with you both!

I am also extremely grateful for everything Dr. Jody Culham has done for me that has led me to this point. Four and a half years ago, a course with Jody persuaded me to pursue a degree in psychology and ever since then Jody has afforded me exciting opportunities to learn more and be involved in research – thank you, Jody!

Finally, I would like to thank all of my friends, family, and colleagues who have supported me every step of the way through this thesis and through my undergraduate studies as a whole. Thank you!

Table of Contents

Where Do I Know That? A Distributed Multimodal Model of Semantic Knowledge

Which is harder to name, a pencil or a banana? For most, neither is particularly difficult to identify, but this is not always the case. There have been many reports of focal neurological damage (e.g., damage resulting from herpes simplex encephalitis) selectively impairing one's ability to identify either living kinds, like bananas, or nonliving kinds, like pencils (Blundo, Ricci, & Miller, 2006; Warrington & Shallice, 1984). The observance of these so-called category-specific deficits following focal brain damage suggests that semantic knowledge has organized, rather than random, mapping in cortex. Were semantic knowledge mapped randomly, focal damage would be expected to produce more random errors instead of the observed systematic (i.e., category-specific) errors. This thesis presents a novel connectionist model of semantic knowledge, which expands on the hub and spoke type model suggested by Patterson, Nestor, and Timothy (2007), to further investigate the structure of semantic knowledge through simulated priming and lesioning in a picture naming task.

Development of the Semantic Feature Model

In 1972, Endel Tulving formally coined the term "semantic knowledge", describing it as "knowledge of the world around us" (Tulving, 1972). Tulving further describes semantic knowledge as being explicit knowledge which can be acquired through either perception or instruction (e.g., by reading). However, it was Collins and Quillian (1969) who questioned the structure of such knowledge and did so prior to Tulving's coining of the name.

Early Semantic Structure. Semantic networks, a conceptual framework pioneered by Collins and Quillian (1969), are hierarchical systems in which concepts (e.g., dog) are highly categorized and features (e.g., <has 4 legs>) are stored at the highest level in which they may remain true. For example, canary is under the domain of bird, which is under the domain of

animal. Features of canaries that are true of all animals (e.g., <eats>) are stored at the animal level, features of canaries that are true of all birds (e.g., <has wings>) are stored at the bird level, and features of canaries that are not shared with all other birds (e.g., <is yellow>) are stored at the canary level. Semantic networks offered a good starting point for the research of semantic knowledge structure, but this perspective has three fundamental flaws. The first flaw is the need to reorganize the system frequently as new information is obtained. If one learned of a wingless bird, <has wings> would have to be moved to the level of each bird that has wings. The second flaw is in the gross (and biologically implausible) redundancy. For example, <is yellow> must be duplicated for each yellow bird. The final shortcoming of semantic networks, and a major reason that prototype theory later dominated, is its inability to explain why research participants were faster to confirm that canaries are birds than they were to confirm that ostriches are birds despite canary and ostrich being equidistant from the bird tier (Rosch, 1973; 1978).

Contributions of Prototype Theory. Prototype theory was first proposed in 1973 by Eleanor Rosch and can generally be considered an extension of the semantic networks approach. In prototype theory, categorization is somewhat preserved (e.g., canary is still considered a subcomponent of bird), but categories are made less concrete as they are defined by a prototype containing both essential features as well as common features (Rosch, 1973; 1978). In birds, <has wings> would be an essential component of the bird prototype while <can fly> would be a common component because all birds have wings but not all birds can fly. It follows that a canary is more prototypical of the bird category than is an ostrich (i.e., canary has more of the typical bird features without sacrificing essential features), and that this difference in prototypicality may drive the observed difference in reaction times to confirm that each is a bird. While Rosch never explicitly stated so, this logic suggests that concepts are merely a collection

of features, which is the core premise of the semantic feature model described next (Smith, Shoben, & Rips, 1974).

Semantic Feature Model. Smith and colleagues were the first to formally propose that concepts are made up of a collection of feature activations (Smith, et al., 1974). In this model, categorizations like bird are largely an emergent property of feature sharing rather than an organizational pillar. Additionally, only one copy of each feature is required because all concepts can activate or not activate a shared feature set. For example, canary and ostrich both activate <has wings>, but only canary activates <can fly>. This reduction makes the model more biologically plausible as biological systems tend not to involve high levels of redundancy. The present model, as well as many other connectionist models which started to emerge in 1976, represent concrete concepts (e.g., cow) as simply a pattern of basic-level feature activations (e.g., <has 4 legs>, <moos>, and <eats> are activated while <quacks>, <used for drawing>, and <flies> are not; Anderson, 1976).

Theories on the Distribution of Semantic Knowledge

There are two primary theories on the distribution of semantic knowledge. The first, domain-specific theory, posits a topography based on an animate-inanimate distinction wherein a region for each of the two kinds has evolved to specialize in differentiation within that kind (Caramazza & Shelton, 1998). In this topography, category-specific deficits would occur when regions of the brain which encode a category of object are damaged. For example, damage to cortex specialized in distinguishing between living kinds would produce a preferential or exclusive deficit in one's ability to name living kinds relative to nonliving kinds.

The second theory, sensory-functional theory, posits that sensory knowledge (e.g., <is red>) and functional knowledge (e.g., <used to open cans>) are stored separately and that the

animate-inanimate distinction is a consequence of differential reliance on sensory and functional knowledge types (Warrington & Shallice, 1984). Damage to functional knowledge would selectively impair nonliving kinds because these kinds most often have a function whereas living kinds typically do not (e.g., a wrench has a function while a lemur does not). Additionally, damage to sensory knowledge would preferentially impair living kinds because, while both kinds have visual features, identifying living kinds depends near-exclusively on visual cues due to their lack of associated functional information.

Goldberg, Perfetti, and Schneider (2006b) had participants perform an association task involving perceptual knowledge retrieval in a functional magnetic resonance imaging (fMRI) machine in order to investigate the distribution of sensory semantic knowledge across cortex. The authors found that retrieval of each tested modality (tactile, gustatory, auditory, and visual) elicited activity roughly adjacent to the region(s) of cortex currently thought to decode that modality. For example, retrieval of auditory knowledge resulted in increased activity in the left superior temporal sulcus, an area that has been strongly implicated in auditory processing (Seifritz, et al., 2002). This finding supported the hypothesis of sensory-functional theory that semantic knowledge types are individually distributed across cortex. In a second study, Goldberg and colleagues (2006a) had participants perform a similarity-based generation-and-comparison task involving fruits, birds, body parts, and clothing. Rather than activate one of two regions specialized for either living or nonliving differentiation as domain-specific theory suggests would have happened, each category of object produced a pattern of activity spread across multiple regions. For example, fruit names activated medial orbitofrontal (associated with taste and smell processing) and ventral temporal (associated with colour processing) cortex (de Araujo, Rolls, Kringelbach, McGlone, & Phillips, 2003).

While the fMRI findings of Goldberg and colleagues (2006a, 2006b) provided clear evidence for the sensory-functional theory, the debate is not yet over (e.g., see Sartori, Gnoato, Mariani, Prioni, & Lombardi, 2007). However, given the current state of research, the presented model of semantic knowledge was designed in accordance with sensory-functional theory rather than domain-specific theory.

Extending the Dichotomies in Feature and Concept Categorization

Over a span of several years, McRae and colleagues assembled a large-scale set of feature production norms (see McRae, Cree, Seidenberg, & McNorgan, 2005). This set of norms is derived from the responses of approximately 725 individuals and is comprised of 541 concrete concepts associated with 2,526 basic-level features, which makes it the largest of its kind. Analyses of these data supports expansion of both the sensory-functional dichotomy of feature types and of the living-nonliving dichotomy of concept categories (Cree & McRae, 2003).

Expanding Feature Types. Cree and McRae (2003) categorized all features in the feature production norms (later published in 2005) into 10 types of semantic knowledge: visual colour (e.g., <is brown>), visual form and surface (e.g., <has 4 legs>), visual motion (e.g., <walks>), functional (e.g., <used for milk>), olfactory (e.g., <smells bad>), gustatory (e.g., <tastes good>), auditory (e.g., <moos>), tactile (e.g., <feels soft>), taxonomic (e.g., <is a mammal>) and encyclopaedic (e.g., <is domesticated>). For each of these knowledge types (with the exception of taxonomic and encyclopaedic semantics), a neural location could be proposed based on the existing research of knowledge retrieval and so this feature categorization was coined “brain region knowledge-type taxonomy.” Cree and McRae also compared the performance of this taxonomy against Wu and Barsalou’s knowledge type taxonomy (later published, Wu & Barsalou, 2009) using hierarchial cluster analysis of concepts based on

features, which resulted in a nearly identical pattern using substiantially fewer categories and more biologically plausible distinctions (i.e., many of Wu and Barsalou's knoweldge types cannot currently be mapped to a particular brain area). Goldberg, Perfetti, Fiez, and Schneider (2007) found imaging evidence that encyclopaedic knowledge is located in left prefrontal cortex, leaving only taxonomic knowledge without a proposed neural location. Consequently, the present model impliments all feature modalities proposed in the brain region knowledge-type taxonomy except for taxonomic semantics.

Expanding Concept Categories. In addition to investigating feature types, Cree and McRae (2003) used the feature production norms to research categorizations in the concepts. Specifically, multiple-factor hierarchical cluster analysis was performed using the brain region knowledge-type taxonomy, feature informativeness, concept confusability, visual complexity, familiarity, and name frequency. Four main concept categories were revealed; living kinds expanded to a) animals and b) fruits and vegetables, while nonliving kinds expanded to c) artifacts (primarily tools, utensils, containers, and clothing), and d) musical instruments. Some differences between these categories have been observed in imaging studies (Siri, Kensinger, Cappa, Hood, & Corkin, 2003). Consequently, the present model considers these four categories rather than simply living kinds versus nonliving kinds.

Overview of Connectionist Modeling Fundamentals

Connectionist modeling (also called parallel distributed processing) is an approach to computational modeling that has been championed by McClelland and Rumelhart (see McClelland, et al., 1986; Rumelhart, et al., 1986b) and has proven useful in structure testing and experimental research in neuropsychological systems including semantic knowledge (Farah & McClelland, 1991). In connectionism, higher-order functions such as semantic knowledge

retrieval occurs as the emergent process of a network of simple interconnected units. These units are called nodes and their connections are called projections. Together, nodes and projections form a likeness of neurons in which nodes can either excite or inhibit one another based on the nodes' current activation level (zero to one) and a mutable property of the projection called the weight (negative for inhibitory interaction and positive for excitatory interaction). Nodes with a similar intended purpose (e.g., nodes representing visual colour features such as <is red> and <is green>) are grouped into what is called a pool, which in turn may be grouped into a layer. Nodes may be used for input, output, processing, or any combination thereof. When nodes do not have any interpretable meaning (i.e., not used for either input or output), they are referred to as hidden. The projections between nodes can be unidirectional or bidirectional with separate weights for each projection in each direction. Projection weights are altered during training through a learning algorithm such as back-propagation (see Hintor, 1989; Rumelhart, Hinton, & Williams, 1986) wherein error between the actual output node activations and target activations is calculated and each weight, working backwards, is adjusted such that the network will produce activations slightly closer to the target in the next epoch (set of training trials). Cross-entropy error is commonly used as the error measure in back-propagation because it produces larger error values when an activation is on the wrong side of 0.5 (e.g., larger error when a node is "on" that should be "off" according to the target pattern), but other error measures such as summed square error may be used. The rate at which weights are adjusted during training is called the learning rate. A momentum factor can be used to add a proportion of the prior weight change to the present weight change, in effect speeding learning in the present direction (this factor should be set to zero, no momentum, until the network has begun to learn in an appropriate direction). Changes to the weights may be applied either after each trial within an epoch (called online

learning) or after each entire epoch (called batch learning). In online learning, the system is sensitive to trial order effects. A passive node activation decrease can be applied through a factor called decay. Time is considered in units called time ticks, which represent discrete processing steps. An interval is made up of a definable number of time ticks and roughly represents the amount of time required for activations to spread from one pool to another. For example, a model could have 5 intervals of 6 ticks, resulting in 30 overall time ticks during which activations fully spread between pools roughly 5 times. Implementation of connectionist modeling is greatly simplified through the use of software packages such as PDPTool that have been designed by researchers for researchers (for more information on connectionist modeling, see McClelland, 2013).

Connectionism and Computationalism. The core posit of computationalism is that the mind operates by performing formal operations on symbols. This is highly inconsistent with connectionism, which achieves higher-order function through a plethora of simple, informal operations (emergent properties). Connectionism simply does not incorporate language of thought, which is the main reason some researchers oppose it while others support it. Many well-known researchers including Zenon Pylyshyn, Steven Pinkel, Jerry Fodor, Joel Lachter, Alan Prince, and Thomas Bever have joined in the debate of connectionism versus computationalism with neither side taking clear lead. At this time, support of either side is largely a matter of personal philosophy (for further discussion, see Pinker, et al., 1998).

Connectionist Modeling of Semantic Knowledge

The first connectionist model of semantic knowledge explicitly based around sensory-functional theory was created by Farah and McClelland (1991). This model featured a layer of input/output nodes and a layer of hidden units, which represented semantics. The semantics

nodes were divided into a sensory pool and a functional pool (pools were fully interconnected), with three times as many nodes in the sensory pool, an experimentally determined approximation of real feature type distribution. Individually, the semantic nodes had no interpretable meaning. Arbitrary living and nonliving concepts were generated as random patterns of activations across the semantics layer with nonliving concepts making more use of functional semantics (6.7 versus 2.1 nodes per concept) and living concepts making more use of sensory semantics (16.1 versus 9.4 nodes per concept). Once it was found that the model was able to learn to perform a picture naming task using arbitrary input and output patterns, the authors simulated incremental, randomized damage to each semantic pool individually. Damage to functional semantics resulted in exclusive impairment of nonliving kinds while damage to visual semantics impaired both kinds, but the deficit was greater in living kinds. In both cases, impairments increased in severity with each degree of damage, a phenomenon referred to as graceful degradation. These two patterns of category-specific impairments matched closely with the category-specific impairments observed in patients who had suffered focal brain damage (Warrington & Shallice, 1984). These findings supported the sensory-functional theory account of category-specific deficits, which posits that either living or nonliving kinds may be more impaired depending on the location of focal damage because each kind relies differentially on sensory and functional semantics, which are located separate from one another and so may be damaged independently.

One of the first connectionist models of semantic knowledge to make direct use of feature production norms was Cree, McRae, and McNorgan (1999; using norms from McRae, de Sa, & Seidenberg, 1997), which simulated semantic priming using three simple layers: a) wordform input, b) semantics output, and c) hidden semantic structure. By using empirically derived concept-feature relations, the authors produced highly interpretable and generalizable results

which ultimately supported the theory that semantic-similarity priming occurs as a result of featural overlap.

Semantic Dementia and a So-Called Semantic Hub

In recent years, connectionist models of semantic knowledge have begun to include divisions of semantics into pools based on knowledge-type to allow each modality to function more independently (i.e., multimodal models). For example, Rogers, Lambon, Garrard, Bozeat, McClelland, Hodges, and Patterson (2004) divided semantics into names, perceptual features, functional features, encyclopaedic features, and visual features. Additionally, the authors included a cross-modal convergence zone through which the semantic modalities could form associations and interact. The existence of such a convergence zone, now called the "hub", in the brain was first proposed in 1989 and has been suggested to reside in anterior temporal lobe (see Damasio, 1989; Patterson, et al., 2007). In the context of connectionist modeling, the hub has been implemented as a pool of hidden nodes which do not represent features and are highly connected with most or all pools of features. However, the existence of this hub remains controversial due to inconsistent imaging findings, partially because anterior temporal lobe is a particularly difficult region to record with fMRI (Patterson, et al., 2007). In connectionist models, convergence hubs have been shown to increase a models ability to form associations, and simulated damage to this hub has been associated with impairments similar to semantic dementia (a global, progressive, and severe deficit to most or all tasks requiring semantic knowledge often associated with neurological damage to the anterior temporal lobe, especially left), however, these effects are still under investigation (Rogers, et al., 2004; Rogers & McClelland, 2004). In rare cases, semantic dementia has been associated with category-specific deficits (Barbarotto, Capitani, Spinnler, & Trivelli, 1995; Lambon Ralph, Howard, Nightingale,

& Ellis, 1998; Lambon Ralph, Patterson, Garrard, & Hodges, 2003; Zannino, et al., 2006), most commonly involving preferential impairment of living kinds (for discussion, see Merck, et al., 2013).

Model Architecture and Preparations

This thesis presents a large-scale, distributed, multimodal connectionist model of semantic knowledge designed to mimic ventral stream processing in order to simulate picture naming (see Figure 1 or Appendix A). The model includes 13 pools containing a total of 2,840 nodes connected through 2,286,180 weighted projections, and was scripted using the PDPTool v2.06 platform with no major modifications (see recurrent back-propagation through time in McClelland, 2013).

Semantic representations of concrete concepts across basic-level features were derived from feature production norms (McRae, et al., 2005), and were divided into eight relatively sparce modality-specific pools: visual colour (34 nodes), visual form and surface (544 nodes), visual motion (95 nodes), functional (794 nodes), olfactory-gustatory (19 nodes), auditory (55 nodes), tactile (39 nodes), and encyclopaedic (739 nodes). In each semantics pool, nodes precisely corresponded with a single feature in the norms (e.g., a node for <is red> resides in the visual colour pool). As was previously stated, taxonomic features were not included. Additionally, the two chemical sensory modalities, gustation and olfaction, were combined due to their low feature counts and similar evidenced locations in cortex (Goldberg, et al., 2006a).

In addition to modality-specific semantics pools, the model includes a hub (150 nodes) through which the modalities may interact and associations may form. In the interest of interpretability, the present model does not have any direct projections between semantic modalities, despite evidence in fMRI literature for such tracts (e.g., between auditory and visual

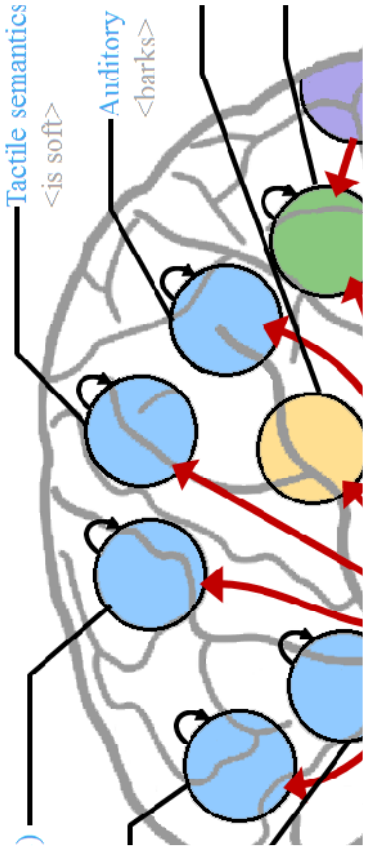


Figure 1. Architecture of the presented model. All arrows indicate full interconnection in the direction(s) indicated. The number of nodes in each pool is displayed in parentheses. Visual semantics are shown in green, nonvisual semantics in blue, wordforms in beige, the hub in red, and a visual decoding layer is shown in purple. Examples of each semantic for the concept of "dog" are indicated in grey below each label. Pools are displayed over the regions of cortex which have been implicated in their processing in prior studies (positions are approximate and sizes are not to scale).

cortex; Beer, Plank, & Greenlee, 2013). Future research may investigate the consequences of introducing direct intermodal connections.

The hub was also directly connected to a wordform pool (30 nodes). Each concept was assigned a random represenation using 3 of the 30 nodes (e.g., <00001 00000 00000 00010 00000 00100>). Use of wordform nodes per concept had aproximtely normal distribution (M = 54.1, SD = 6.21, $Range$ = 41 to 66). Note that this implimentation of wordforms is not intended to replicate true wordform processing and may instead be considered as an abstraction of phonology and/or orthography. Wordform nodes were used exclusively for output in this implimentation, but could be used for input in the future.

In order to reproduce picture naiming, a pool for visual inputs (240 nodes) was included in the model along with a smaller pool (100 nodes) to decode this input before it reached visual semantics. The input patterns for this pool were created from line drawings of the 541 concepts in the feature production norms in an effort to somewhat preserve visual properties such as outline shape and visual complexity. Drawings were primarily obtained from Bonin and colleages (2002), Rossion and Pourtois (2004), and Szekely and colleages (2005). Images not available from academic sources were obtained using google image search with settings set to return only images with usage rights labeled for free reuse and modification. All images were binarized to black and white (“on” and “off”), cropped to center, and resized to exactly 150 by 150 pixels (see Appendix B). The resulting 22,500 pixels in each image would have been unreasonably large input so a reduction process refined these to 240 units (see Figure 2) as follows. First, pixel usage was quantified by generating a heat map from all images, which revealed that pixel usage was highest in the center and decreased gradually towards the perimeters. Second, the 150 by 150 area was sectioned off into a 5 by 5 grid. Third, the most

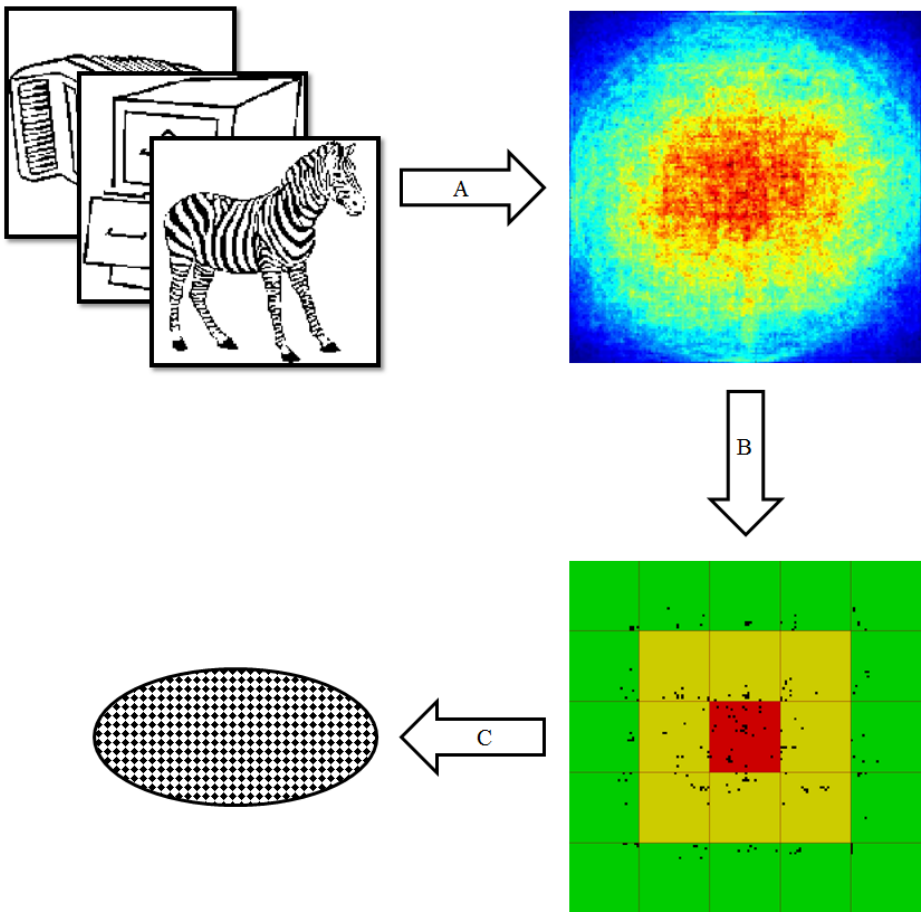


Figure 2. Preparation of visual input from line drawings. (A) Once all 541 images were binarized, centered, and resized, they were summed to determine overall pixel variability and placement. (B) The most variable pixels from each cell were noted (black), selecting 30 pixels from the center (red), 10 pixels from each middle cell (yellow), and 5 pixels from each outer cell (green). (C) The noted pixels' values were extracted from each image to produce a pattern across 190 binary units for each image. 50 additional units were added to serve as padding for images that had minimal representation in the selected pixels. All images with fewer than 20 of the selected pixels had random padding units added to their representation until it contained 20 units. The result was a pattern containing at least 20 of 240 units for each of the 541 images.

variable pixels from each cell were selected, choosing 30 pixels from the center-most cell, 10 from each midle cell, and 5 from each outer cell. This method ensured that generally informative pixels (i.e., those which were used frequently) were selected without sacrificing any region of the images, and was appealing due to its similarity to foveal-perifoveal-peripheral visual acuity in the eye. Fourth, the pixel values (on or off) in the 190 selected locations were extracted from all images. In roughly one quarter of concepts' images, fewer than 20 pixels in the selection were "on" so 50 padding units were added (to the 190 units) and these concepts had a random selection of padding added to their representation such that all represenations had at least 20 of 240 units "on" (see Figure 3). Last, the visual representations were correlated with one another to confirm that no two represenations were overly confusable (see Figure 4). The use of these representations to activate all three visual modalities (visual colour, motion, and form) is justified by the abstraction in their preparation as well as by the use of a visual decoding pool. If the visual inputs only activated form and surface information (as one might expect line drawings to do), then the form and surface pool could not be meaningfully lesioned as doing so would completely "blind" the network, making picture naming impossible. Using all three visual semantics as input relays to the hub makes the system relatively less vulnerable to visual semantic damage, thereby allowing these pools to be appropriate targets in lesion experimentation.

The final pool included was a bias pool (1 node), which is not shown in any figure. The bias pool (node) was always fully active and had projections to all nodes except those in the hub, visual decoding pool, and visual input pool. These projections were all unidirecitonal and trainable. The role of the bias pool is to systematically control the likelihood that each unit is "on" or "off" regardless of input, in effect altering the y-intercept of each connected node.

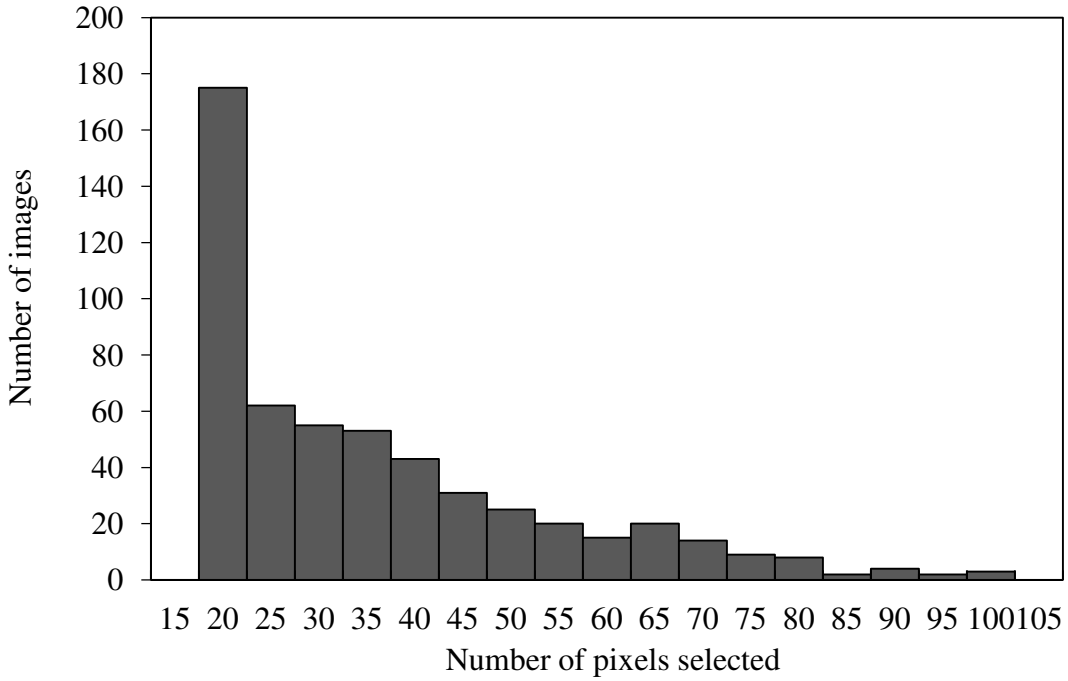


Figure 3. Number of selected pixels present in each image. Note that padding units are included.

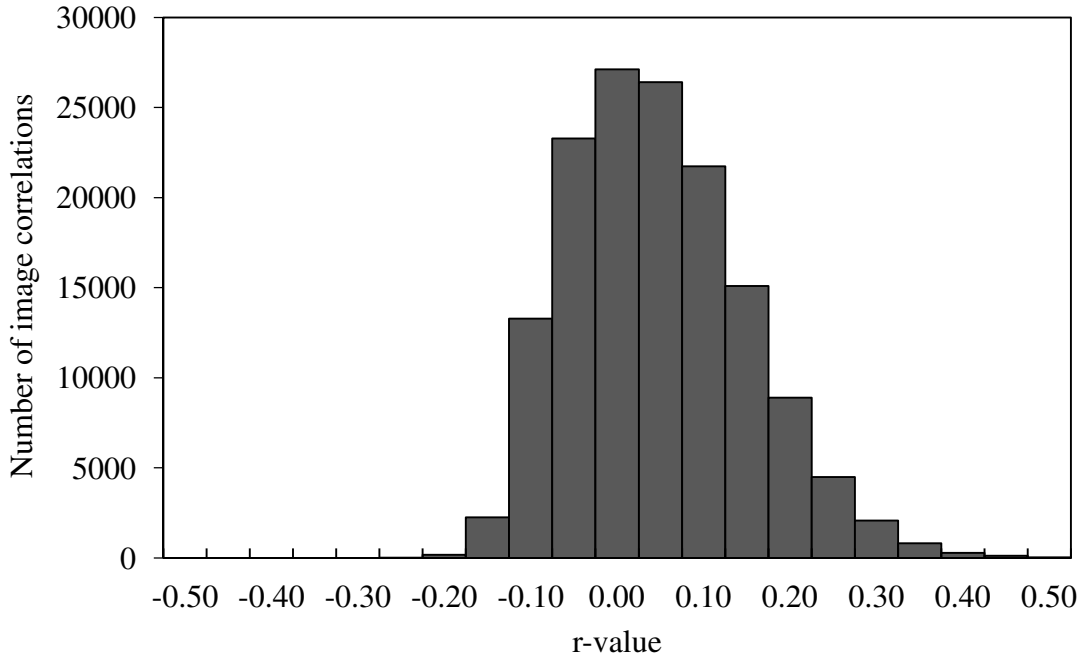


Figure 4. Correlation values (r-values) between all pairings of abstracted visual representations produced from line drawings. On average, the representations were very minimally correlated

Including a bias pool assists model learning and is included by default in PDPTool (McClelland, 2013). In this model, the bias pool served to keep semantic and wordform nodes "off" as a default.

The spread of activations in the present model was designed to be as follows. First, visual input activates visual decoding nodes through unidirectional full interconnection. Next, visual decoding activates visual colour, motion, and form semantics through unidirectional full interconnection. Within each visual semantic pool, activations can spread between nodes through full bidirectional interconnections (e.g., the model will directly increase the activation in the node for <walks> when <runs> is active once it learns that things which run can also walk; nonvisual semantic pools also have these intramodal projections). As visual nodes interact, activation also spreads into the hub via full bidirectional interconnections. Hub nodes do not directly influence other hub nodes, but are fully and bidirectionally connected to every semantic node (visual and nonvisual) and wordform node. Consequently, activation can spread to all pools (except visual input and decoding) once the hub is reached. At this stage, a frenzy of activations (hub to semantics), co-activations (features within modalities), and interactions (features between modalities through the hub) occurs, eventually settling on a final pattern of activations.

Model Training

The model was trained once for all subsequent experiments. This training involved two tasks, both of which occurred over six intervals of five time ticks (30 time ticks in total). The first task was picture naming and feature production, during which visual inputs were supplied (hard-clamped for the entire duration) and the targets included the wordforms and all semantics. The second task was concept naming from features, during which inputs were given to all

semantics (hard-clamped for the entire duration) with only wordforms as a target. The first task prepared the model for picture naming while the second task ensured that nonvisual semantics would form associations with visual semantics and also contribute to wordform activations. During each epoch (training cycle), each of the 541 concepts were trained on both tasks in a fully randomized order for a total of 1,082 trials per epoch. Weights were updated according to back-propagation after each trial (online learning) based on the cumulative cross-entropy error calculated across the final 10 time ticks, defined as:

$$-1 * \frac{\sum_{i=21}^{30} \sum_{n}^{N} t_{ni} \ln(a_{ni}) + (1 - t_{ni}) \ln(1 - a_{ni})}{10}$$

wherein *i* is the time tick, *n* is the node, *N* is the number of outputting nodes across all pools, t_{ni} is the target (intended) value for node *n* during tick *i*, and a_{ni} is the actual activation in node *n* during tick *i.* Learning rate was 0.005 and no decay factor was applied. Momentum was held at 0 until the 11th epoch, at which time momentum was increased to 0.9 to speed learning. Before training, all weights were initialized to random values between -0.05 and +0.05. Before each trial, all activations were zeroed so that activations from the prior trial did not affect the present trial except through changes to the projection weights.

Initially, the model was trained until at least 95% of the nodes intended to be "on" were active greater than 0.7 after each trial and at least 95% of those intended to be "off" were active less than 0.3 after each trial. At this time (epoch 53), it became apparent that visual semantics learn to activate much more slowly than do nonvisual semantics during the picture naming and feature production task in this architecture. By the time visual semantics reached the 0.7

activation threshold, nonvisual semantics already had means in the extreme 0.9s. To prevent giving nonvisual semantics an unjustified advantage over visual semantics, training was continued until all final target activations were within a range of 0.01, which occurred after epoch 83. At this time, all mean target activations were approximately 0.99. The binary nature of activations in visual semantics (the only gateway to the hub during picture naming) in conjuncture with the lack of intrapool projections in the hub (would have facilitated more complex and potentially problematic processing) makes it unlikely that overtraining was an issue. For the time course of average activations in the trained network, see Figure 5.

Introducing Error and Accuracy Measures

Evaluation of nearly all testing of the proposed model involves a cross-entropy error measure that has been standardized for pool size (referred to as PCEE), which is based on final activations and was defined as:

$$PCEE_{pool} = -1 * \frac{\sum_{n}^{N} t_{nf} \ln(a_{nf}) + (1 - t_{nf}) \ln(1 - a_{nf})}{N}$$

wherein n is the node, N is the number of nodes in the pool, t_{nf} was node n's final target, and a_{nf} was the node n's final activation. This error measure reflects the degree to which a pool's node activations and non-activations are incorrect with larger PCEE indicating less correct activations. The PCEE serves as a simple single measure that can be used for comparing the activation correctness of two or more pools of any size.

In addition to PCEE, a naming accuracy (NA) measure is used to evaluate overall performance. Following each trial, the final activations in wordform nodes are compared against all 541 wordform representations using normal cross-entropy error. If the correct representation

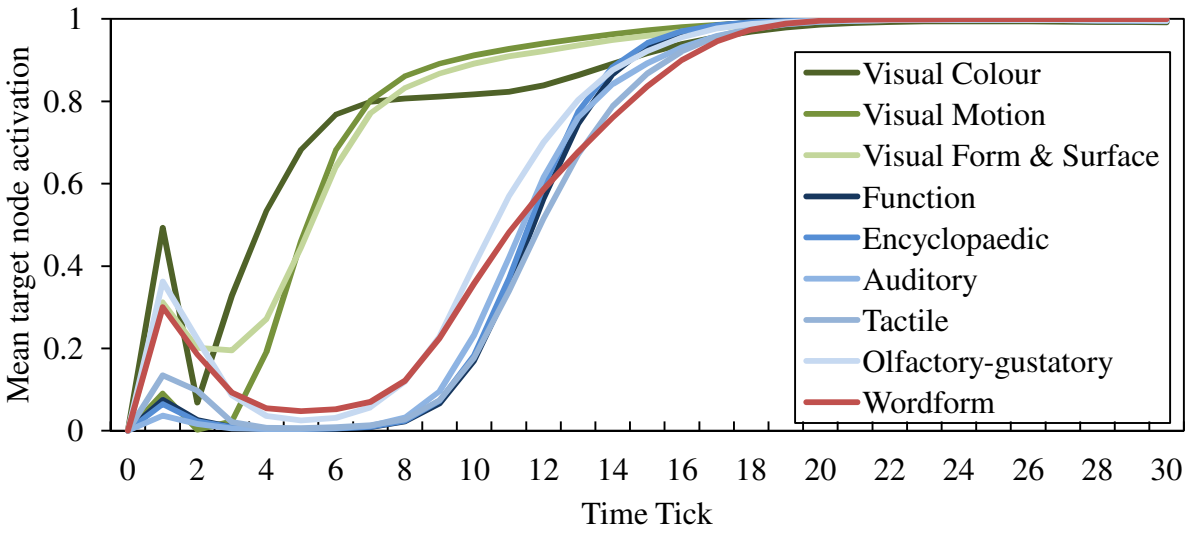


Figure 5. Time course of target node activations of an average picture naming trial after training. The mean activation over time is displayed for an average node that is intended to turn "on" (approach a value of 1). Visual semantics come online first as a result of the input modality. Wordform activation begins with nonvisual semantics, but is the last to plateau.

produces the smallest error (i.e., smaller than for any of the other 540 representations), then the trial is deemed a success. The NA represents the percentage of trials which were deemed successful. For example, a NA of 60% means that the correct wordform pattern produced the smallest error in 325 of the 541 concepts (approximately). This offers a highly interpretable measure of overall performance with regard to the picture naming task without using arbitrary cut-off values. After training, the network was able to achieve 100% NA.

When neither PCEE nor NA are appropriate, the mean activation in target and non-target nodes is instead presented. Target nodes should have a mean near one ("on") while non-target nodes should have a mean near zero ("off").

Confirmation of Proper Network Function

Each confirmation task uses an adaptation of one of the two training tasks. In all cases, timing was unchanged from training; 6 intervals of 5 ticks and all specified inputs were hard-clamped for the entire duration. Pools were disconnected from the network by zeroing all weights to and from nodes in pool as specified.

Images activated visual semantic modalities. First, it was important to confirm that image inputs correctly activate visual semantics (colour, motion and form). This was achieved by running the picture naming task used in training for all 541 concepts. The final activation value for each node was recorded after each trial. Across all concepts, target nodes averaged a final activation >0.99 for each of visual colour, motion, and form, and non-target nodes averaged a final value <0.01 for each visual semantic modality. As expected, visual semantics were correctly activated or not activated.

Visual semantic modalities used the hub to co-activate. To determine whether visual semantic modalities were successfully using the hub to interact with other visual modalities,

visual inputs were provided and final activations were recorded in only visual semantics. All trials were repeated twice as follows. In the first repetition, the visual modalities were connected to the hub, but the hub was cut-off from all nonvisual semantics and wordforms. In the second repetition, the visual modalities did not have any access to the hub. In this reduced system, removal of the hub increased PCEE in visual modality pools by 16 to 30 times, which implies that visual modalities were interacting through the hub to help activate one another.

Activation of visual semantic modalities was modulated by nonvisual semantics. To determine if activation of visual modalities was assisted by nonvisual modalities, visual input was provided with and without access to nonvisual modalities and PCEE was calculated for visual modality pools. Removal of nonvisual semantics increased PCEE in visual semantics by 35 to 55 times, indicating that nonvisual semantics are relevant to the activation of visual semantics.

Visual semantic modalities tap into nonvisual semantics. The ability of visual semantic modalities to independently activate wordforms was investigated by removing all nonvisual modality pools. Removal of the nonvisual semantics resulted in a NA decrease from 100% to 68.39%, meaning visual semantics alone are sufficient for moderate accuracy, but access to nonvisual semantics is necessary for full accuracy.

Nonvisual semantic modalities tap into visual semantics. The ability of nonvisual semantic modalities to independently activate wordforms was investigated by applying input directly to nonvisual semantics with and without access to visual semantics. Loss of access to visual semantics decreased NA from 100% to 53.03%. Similar to visual semantics, the nonvisual semantics contributed to correct picture naming, but required full access to all semantics to achieve high accuracy.

All semantic modalities tap into wordform knowledge. Lastly, the roll of wordforms in visual and nonvisual modality activation was access by severing the connection between the hub and wordforms. Without access to wordforms, PCEE increased by 16 to 26 times in visual semantics pools and 15 to 50 times in nonvisual semantics pools. Olfactory-gustatory, auditory, tactile, and form and surface semantics were particularly affected. These findings indicate that, in addition to activating the correct wordforms, the network uses these wordform activations to further activate all semantic modalities.

Overall, these simulations confirm that, in this model, a) picturing naming taps into nonvisual semantics as intended, and b) the hub was used to form extensive associations between semantic modalities as well as wordforms.

Experiment 1: Replication of Cree, McRae, and McNorgan (1999)

To assess the validity of the presented model, the third simulation in the previously described Cree and colleagues (1999) priming study was replicated. In the original study, a series of concepts were primed with featurally similar and dissimilar concepts using a simple three layer model informed by feature production norms. The authors presented the wordforms of the prime for three intervals of five ticks before presenting the target wordform. The cross-entropy error in semantics was then measured for 20 time ticks as the activations shifted and settled for the new target. Priming effects (lower error measures at all time points) were observed for featurally similar parings regardless of priming order (see Figure 6). For further details including the concept lists, please see the original publication.

Method. The present model was trained to simulate picture naming, not feature production from names, so changes had to be made to the original paradigm. Prime images were presented for three intervals of five ticks after which the target images were presented and cross-

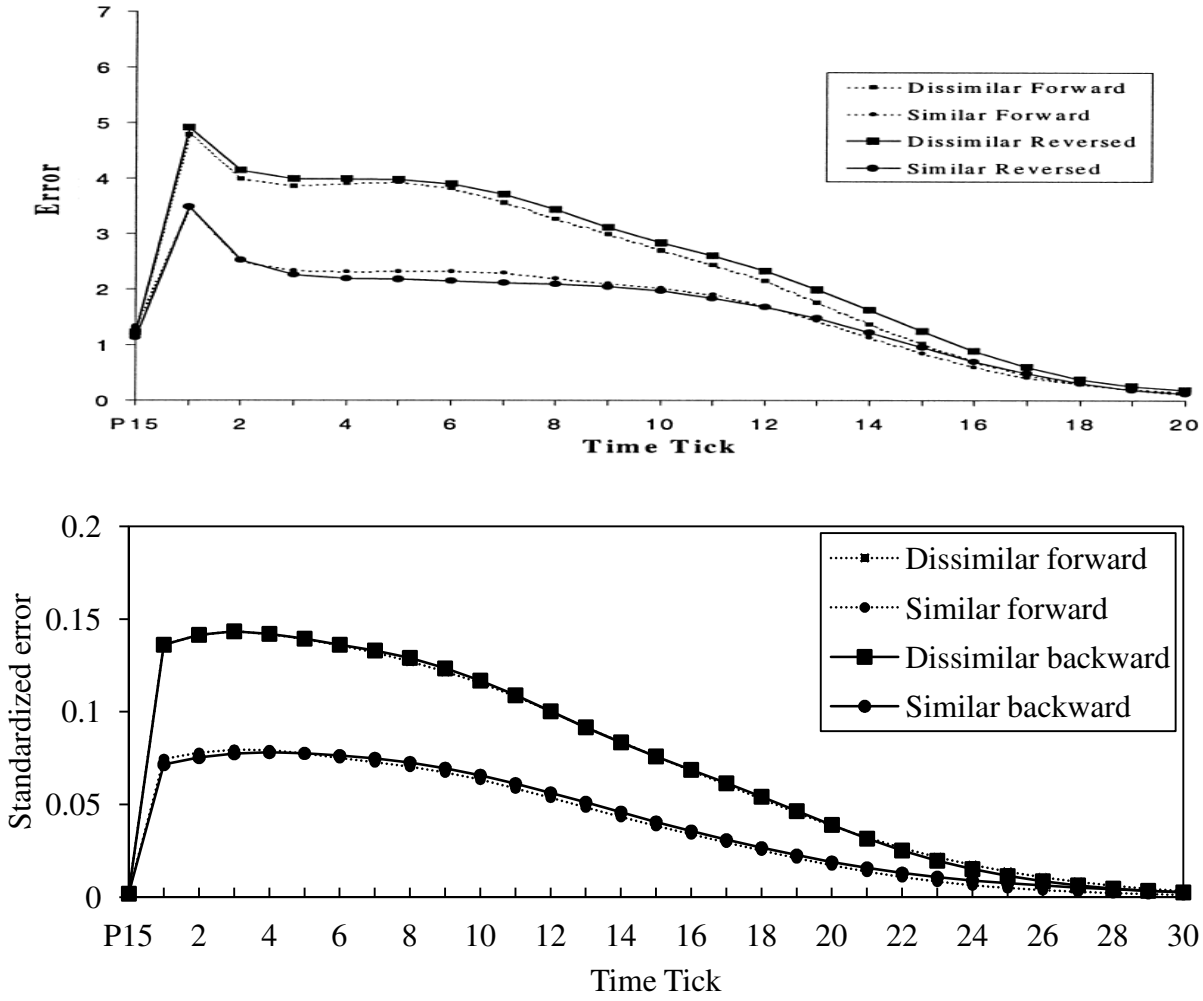


Figure 6. Semantic error across time following presentation of a prime. (Upper) Figure 5 from Cree et al (1999), adapted with permission. (Lower) Standardized cross-entropy error in combined semantics and wordforms across time following 15 ticks of visual semantic priming. In both plots, the x-intercept represents the final priming tick (P15) and the error indicated therein was calculated for the prime rather than the target. Both networks produced a priming effect dependent on featural similarity.

entropy error (standardized for number of nodes) was calculated across all semantic and wordform nodes. It should be noted that the meaning of an interval depends on the structure (i.e., number of layers) and so even the timing could not be fully mimicked. Additionally, the error measures used differ slightly (mostly in the meaning of magnitude). Furthermore, settling of semantic activity in the present model required 30 ticks on average rather than the 20 ticks reported by Cree and colleagues. This will have been partly due to the additional layers that activations must spread through in the present model.

Results and discussion. The present model produced a similar pattern of error over time and thereby reproduced the featural similarity-based priming effect observed in the original study (see Figure 6). This successful replication serves to affirm the validity of the present model.

Experiment 2: Simulating Damage in Visual and Functional Semantics

Many prior models dating back to Farah and McClelland (1991) have focused on a visual-functional semantics distinction with regard to category-specific deficits. In this experiment, two progressive lesions were simulated; one covering all three visual semantics types and the other focal to the one functional semantics type. The effects of these lesions were compared across the previously stated categories; animals, artifacts (tools, utensils, containers, and clothing), musical instruments, and fruits and vegetables (as described in Cree, et al., 2003). Together, these categories span 306 of the 541 concepts. The presence of each semantic modality in each category is summarized in Table 1. Based on previously stated findings, it was hypothesized that functional damage would result in an exclusive impairment of nonliving kinds (primarily artifacts) and that visual semantics damage would result in a preferential impairment of living kinds (primarily animals). High degrees of visual semantics damage will simulate a "semantic blindness", making all picture naming impossible.

Table 1.

Average number of features per concept by feature type and concept category.

Modality	Animals (133)	Fruits and Vegetables (55)	Artifacts (100)	Musical Instruments (18)
Encyclopaedic	3.17	3.00	1.73	0.94
Function	0.92	2.96	3.94	2.33
Gustatory-olfactory	0.08	1.31	0.03	0.00
Auditory	0.45	0.00	0.03	1.83
Tactile	0.33	0.80	0.58	0.06
Visual colour	1.02	1.78	0.67	0.67
Visual form and surface	4.52	3.20	4.89	3.78
Visual motion	2.04	0.00	0.01	0.00

Note: The number of concepts in each category is expressed in parentheses.

Method and predictions. For each planned lesion, projections were selected from those directed in to, out of, and within each pool involved. Projections from visual decoding and from bias were never targeted. Selected projections had their weights set to zero to simulate complete knockout. The damage degrees 0%, 25%, 50%, 75%, and 100% damage were simulated. For 25%, 50%, and 75%, selection of projections was random within the targeted projection population. Consequently, it was essential that these damage degrees be repeated (25 repetitions each) and averaged to determine the typical effect. The NA measure was used to evaluate results.

Results and discussion. See Figure 7. As predicted, functional semantic damage yielded impairments primarily in artifacts. Musical instruments were the next most affected followed by fruits and vegetables. Animals were very minutely affected. Overall, being animate and/or biological predicted impairment by functional damage with the further distinction that artifacts were more impaired than musical instruments (both of which are of course inanimate and non-biological). Interestingly, functional semantic use did not fully predict these effects as musical instruments had fewer functional features on average than do fruits and vegetables, but their naming was more affected by damage to these features' nodes.

The co-occurrence of semantic blindness with overall visual semantics damage complicates its interpretation, but moderate (25% and 50%) collective visual semantics damage did impact animals most and artifacts least. Experiment three further investigated these effects of visual semantic damage.

Experiment 3: Simulating Visual Modality- Focal Semantics Damage

Method and predictions. Each visual semantic modality (visual colour, visual form, and visual motion) was individually impaired using the categories and procedures outlined in Experiment 2. Visual motion semantics damage was expected to exclusively impair animal

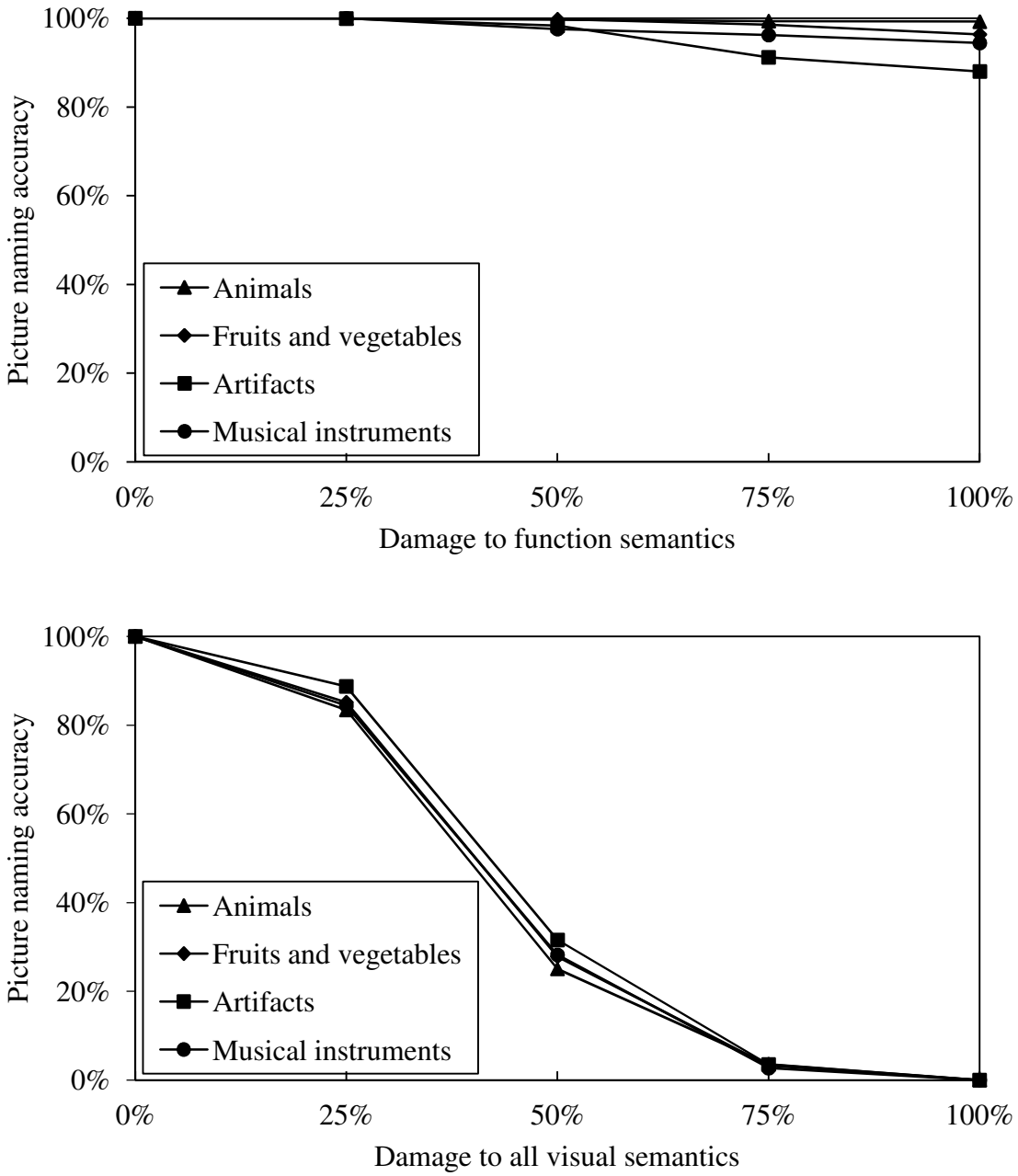


Figure 7. Picture naming accuracy following simulated general visual and nonvisual semantics damage. (Upper) Naming accuracy following incremental damage to functional semantics shows preferential impairment of artifact concepts. (Lower) Naming accuracy following incremental damage to all visual semantics: colour, form, and motion, shows slight relative preservation of artifacts prior to severe semantic "blindness".

naming because only the animal category made use of these motion semantics (the only exception was the artifact “cape” which <flows in the wind>). Visual colour semantic damage (a likeness to cerebral achromatopsia) was expected to preferentially impair naming of fruits and vegetables because these concepts contained more colour features than any other. Additionally, many prior studies have found strong associations between colour processing and fruits and vegetables (Samson & Pillon, 2003). Form and surface information is supremely important for identifying all categories and these features were highly utilized by all. Patient data involving knowledge deficits in the form and surface semantics, as described by Cree and McRae (2003), could not be found. Consequently, the only prediction for visual form and surface semantics damage was that impairments would be large, widespread, and might follow the pattern of mean visual form features per concept (i.e., artifacts and animals are predicted to be impaired more than musical instruments and fruits and vegetables, refer to Table 1).

Results and discussion. See Figure 8. Simulated damage to visual colour semantics almost exclusively impaired the naming of fruits and vegetables, as expected. Animals were slightly more impaired than artifacts and musical instruments, which may be explained by the relatively higher number of colour features per animal concept.

As predicted, simulated damage to visual motion semantics exclusively impaired animal naming. The only way in which non-animals could have been influenced by visual motion damage is if the severed projections were required to keep motion features “off”, which could have been the case if the bias pool were not included in this network.

The consequences of visual form and surface semantics damage were unexpected. In addition to the expected severe global impairment, artifacts and musical instruments were particularly impaired, followed by fruits and vegetables. Animals, counter to prediction, were the

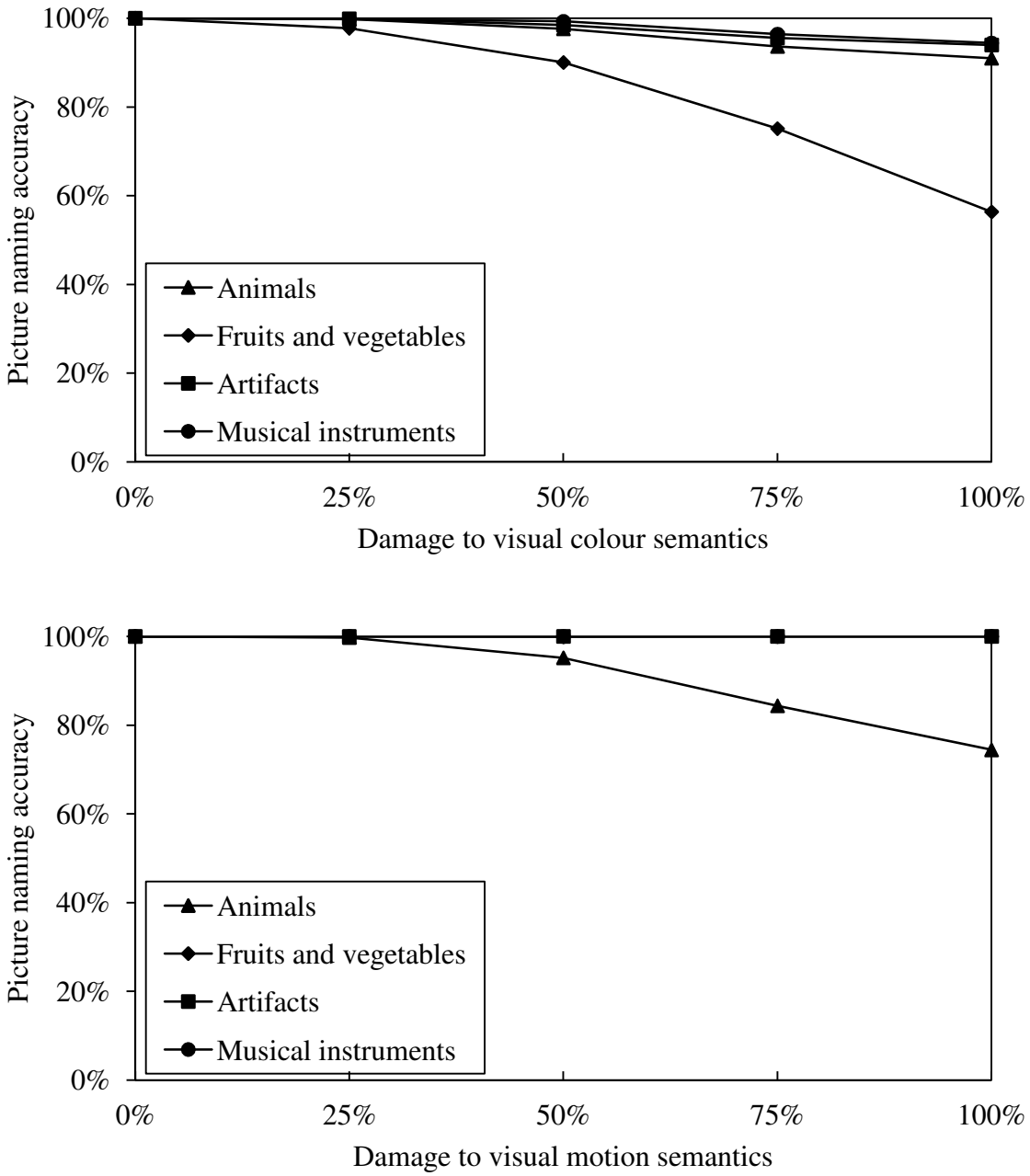


Figure 8. Picture naming accuracy following simulated focal damage to visual semantic types individually. (Upper) Naming accuracy in fruits and vegetables is highly impaired by damage to visual colour semantics. (Lower) Naming accuracy in animals is moderately impaired by damage to visual motion semantics. (Next Page) Naming accuracy is globally affected by visual form and surface semantic damage with preferential impairment of fruits and vegetables over animals and of both nonliving kinds, artifacts and musical instruments, over all others.

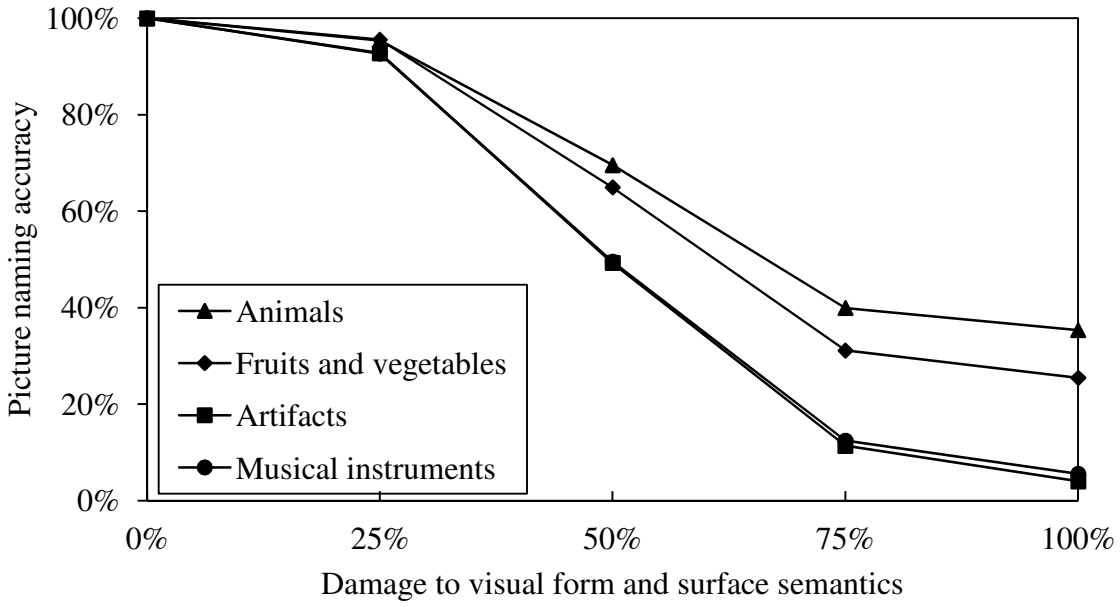


Figure 8. (continued)

least impaired. Perhaps, animals' reliance on visual motion semantics and fruits and vegetables' reliance on visual colour semantics mitigated the effects of form and surface damage. Based on the findings of Cant, Arnott, and Goodale (2009), which evidenced two separate regions for the processing of visual form and visual textures, it is possible that visual form and surface are actually two distinct semantic knowledge types and should have been placed into separate pools. This finding should be subject to further investigation following this thesis.

Experiment 4: Simulating Nonvisual Modality- Focal Semantics Damage

Method and predictions. Following the same method as Experiment 2, nonvisual semantic modalities were individually targeted with simulated focal lesions. The modalities targeted included encyclopaedic, olfactory-gustatory, auditory, and tactile. Recall that functional semantics were already targeted in Experiment 2. Theories on the neural location of encyclopaedic semantic memory are still new (see Goldberg et al., 2007) and so the only predictions made regarding the consequences of damage to these units is that impairment may reflect the number of encyclopaedic features per concept (i.e., animals may be most impaired followed by fruits and vegetables, refer to Table 1). Olfactory-gustatory semantics were utilized near-exclusively by fruits and vegetables. Perhaps unintuitively, very few animals' concepts made use of these chemical sensations. Similar to the case of motion semantics (used only by animals), it was hypothesized that damage to olfactory-gustatory semantics would hinder only the naming of fruits and vegetables.

For both auditory and tactile semantics, no predictions could be generated from patient data and so predictions were made based on feature use. Musical instruments had the most auditory features per concept, followed by animals. Therefore, it was predicted that musical instruments would be most impaired and animals would be the second most impaired. No

impairment was expected in artifacts or fruits and vegetables as neither made notable use of auditory semantics. Similarly, tactile semantics impairments were expected to cause deficits in the naming of fruits and vegetables as well as artifacts, but not animals or musical instruments.

Results and discussion. See Figure 9. Simulated damage in encyclopaedic semantics produced relatively small naming impairments with the greatest impairment in animals followed by fruits and vegetables. Artifacts and musical instruments were not affected. These results matched the prediction and suggest that categories do not generally rely strongly on encyclopaedic information for picture naming.

Simulated lesion in olfactory-gustatory semantics produced no notable deficits in picture naming. It is possible that the preserved colour semantics mitigated a potential fruit and vegetable impairment. A follow-up study should be performed to determine of olfactory-gustatory semantics damage modulates the impact of colour semantics damage on picture naming of fruits and vegetables.

Damage to auditory semantics resulted in a very slight musical instruments naming deficit, but only at 100% damage. This suggests that wordform activations for musical instruments rely mostly on form semantics and functional semantics. It is possible that auditory semantics could produce larger impairments in combination with these other semantics.

Tactile semantic damage was not associated with impairment in any category. This is not overly surprising because tactile features were sparsely present in this concept sample. Tactile semantics are another modality that may act more as a modulator for more impactful damage types (e.g., functional or visual form and surface semantics).

Experiment 5: Simulating Semantic Dementia with Hub Damage

Method The final experiment once again used the same methods described in

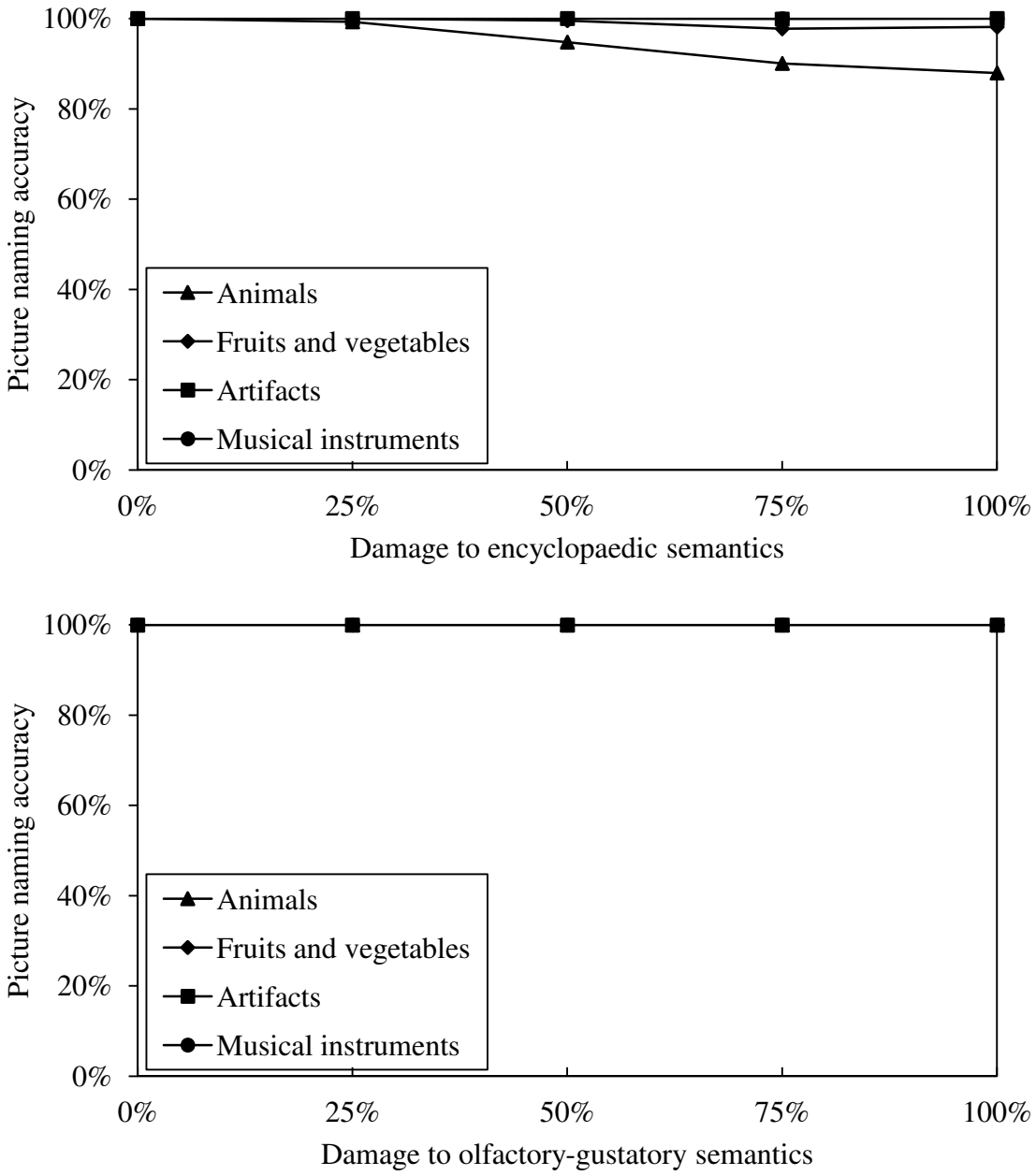


Figure 9. Picture naming accuracy following simulated focal damage to the remaining nonvisual semantic types individually. (Upper) Naming accuracy in animals and to a lesser degree fruits and vegetables was impaired by encyclopaedic semantic damage. (Lower) Naming accuracy was not notably impacted by olfactory-gustatory semantics. (Next Page, Upper) Naming accuracy for musical instruments was slightly impaired at extreme levels of auditory semantic damage. (Next Page, Lower) Damage to tactile semantics did not affect naming accuracy.

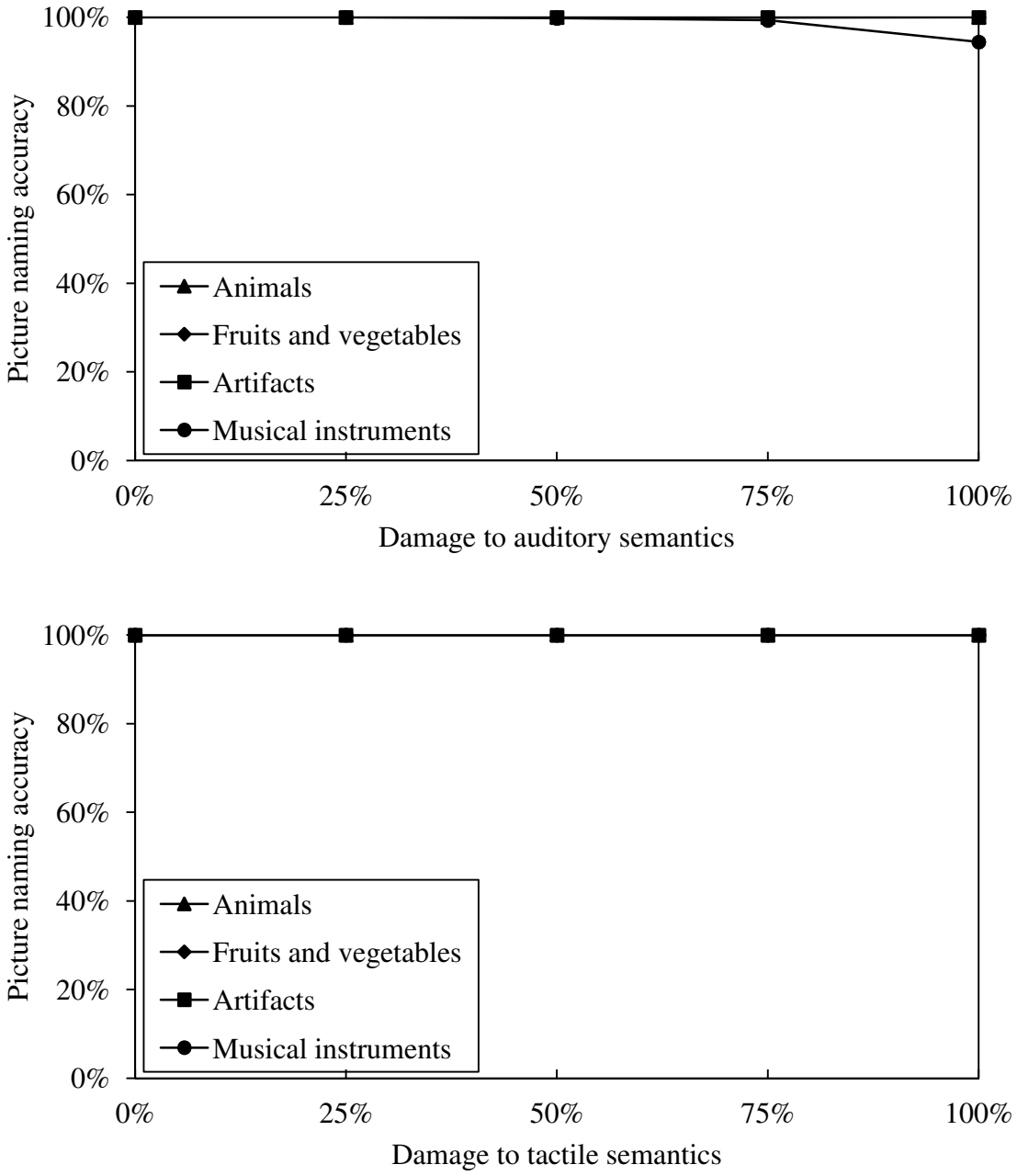


Figure 9. (continued)

Experiment 2 targeting the hub. Specifically, connections in to and out of the hub were damaged. Recall that the hub did not have intrapool projections. In addition to the damage degrees described in Experiment 2, 8.33%, 16.67%, and 33.33% damage were included to increase resolution in the less severe damage degrees. Prior research suggested that one of two results would occur: a) global deficits (i.e., semantic dementia), b) category-specific deficit, most likely impairing living kinds more-so than nonliving (Barbarotto, et al., 1995; Lambon Ralph, et al., 1998; Lambon Ralph, et al., 2003, Patterson, et al., 2007; Zannino, et al., 2006).

Results and discussion. See Figure 10. Unsurprisingly, severe damage to the hub (33.33% and above) completely dissolved the network's ability to name pictures. Prior to 33.33% damage, category-specific deficits were observed with animals most impaired followed by fruits and vegetables, then musical instruments, and least impaired was artifacts. Unlike experiments one through four, this result is not visible at 100% damage, which produces deterministic results that require no statistical evaluation. Consequently, a one-way analysis of variance was performed to evaluate the category-specificity in naming impairments at 25% damage. This analysis yielded a significant effect of category, $F(3,96) = 8.85$ ($p < .001$). To further investigate this effect, post hoc multiple comparisons analysis with Tukey's honestly significant difference criterion was performed, which yielded a significant difference between animals and artifacts only ($p < .05$). This finding suggests that damage to convergence zone(s) alone may be capable of inducing semantic dementia with accompanying category-specific deficits.

General Discussion

So far, only the final node activations have been investigated in the presented model, which ignores potentially rich effects in temporal dynamics. Future research should calculate

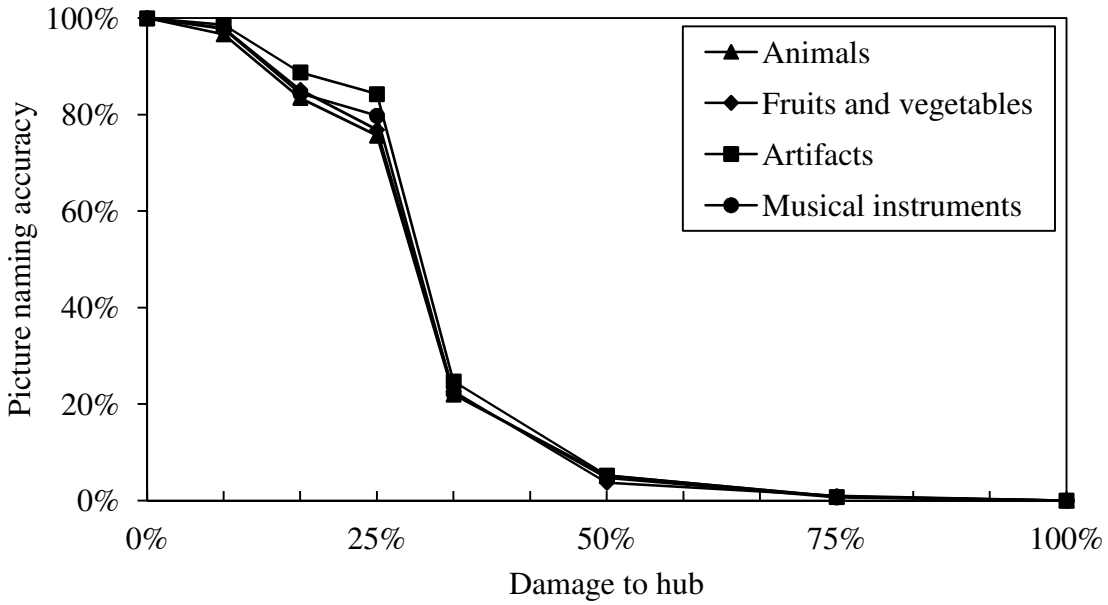


Figure 10. Picture naming accuracy following simulated damage to the hub. At low degrees of hub damage (less than 33%), naming of animals was most impaired followed by fruits and vegetables, then musical instruments, and finally animals. Starting at 33% damage, picture naming was highly inaccurate for all categories.

activation reaction times (first time tick in which wordforms are correct and remain correct) to compare with patient reaction time data. Likewise, the types of naming errors made have not yet been studied in this model. Beyond the NA measure used, proportions of correct names, incorrect names from the correct category, incorrect names from other categories, and non-word responses could be investigated. Non-naming measures should also be developed so that the effects of damage to wordforms can be investigated.

Damage to both tactile semantics and olfactory-gustatory semantics failed to produce any impairment. This could mean that the network was over-trained to a point where base performance is subject to a ceiling effect and substantial damage is therefore required to produce any impairment. Different training tasks may help eliminate this concern.

As previously stated, this model does not include direct intermodal connections which have been evidenced in the human biology. Such connections would alter the role of the hub and perhaps invalidate the present findings on simulated hub damage.

The only task performed was picture naming. Picture naming is a useful test because it is so commonly included in patient studies and (theoretically) taps into all of semantic knowledge. However, picture naming also confounds visual semantic damage effects with a blindness effect that limits the degree to which these visual semantics can be meaningfully studied. Other tasks such as feature production should be tested in future research.

Conclusion

This distributed multimodal model reflects the next advancement in biologically plausible architectures of semantic knowledge. Several of the effects evidenced in prior patient case studies and connectionist model research have been successfully replicated with the presented model despite its complexity. The outcome of experimentation within the presented model

suggests that the category-specific impact of focal lesions may be partially predicted by the degree of damage and the degree to which each category makes use of the impacted semantic modality. The model provides evidence that semantic dementia may be acquired as a result of damage to a convergence zone and further that minor category-specific deficits may be co-acquired. This distributed multimodal model of semantic knowledge performed quite well for a first implementation and certainly merits further development. With proper tuning, this model could be very useful in furthering the understanding of semantic knowledge and its impaired states, and perhaps even aid in the development of treatments for individuals who have sustained certain neurological damages.

References

Anderson, J. (1976). *Language, memory, and thought.* NJ: Lawrence Erlbaum Associates.

Barbarotto, R., Capitani, E., Spinnler, H., & Trivelli, C. (1995). Slowly progressive semantic impairment with category specificity. *Neurocase, 1*, 107–119.

Beer, A., Plank, T., & Greenlee, M. (2013). Diffusion tensor imaging shows white matter tracts between human auditory and visual cortex. *Experimental Brain Research, 213*(2-3), 299-308.

Blundo, C., Ricci, M., & Miller, L. (2006). Category-specific knowledge deficit for animals in a patient with herpes simplex encephalitis. *Cognitive Neuropsychology, 23*(8), 1248-1268.

Bonin, P., Peereman, R., Malardier, N., Méot, A., & Chalard, M. (2003). A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behavior Research Methods, Instruments, & Computers, 35*(1), 158-167.

Cant, J., Arnott, S., & Goodale, M. (2009). fMR-adaptation reveals separate processing regions for the perception of form and texture in the human ventral stream. *Experimental Brain Research, 192*(3), 391-405.

Caramazza, A., & Shelton, J. (1998). Domain-specific knowledge systems in the brain the animate-inanimate distinction. *Journal of Cognitive Neuroscience, 10*(1), 1-34.

Collins, A., & Quillian, M. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8*, 240-247.

Cree, G., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology, 132*(2), 163-201.

Cree, G., McRae, K., & McNorgan, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. *Cognitive Science, 23*(3), 371-414.

Damasio, A. (1989). The brain binds entities and events by multiregional acrivation from convergence zones. *Neural Computation, 1*, 123-132.

de Araujo, I., Rolls, E., Kringelbach, M., McGlone, F., & Phillips, N. (2003). Taste-olfactory convergence, and the representation of the pleasantness of flavour, in the human brain. *European Journal of Neuroscience, 18*(7), 2059-2068.

Farah, M., & McClelland, J. (1991). A computational model of semantic memory impairment: Modality- specificity and emergent category-specificity. *Journal of Experimental Psychology: General, 120*, 339-357.

Goldberg, R., Parfetti, C., & Schneider, W. (2006a). Perceptual knowledge retrieval activates sensory brain regions. *Journal of Neuroscience, 26*(18), 4917-49-21.

Goldberg, R., Perfetti, C., & Schneider, W. (2006b). Distinct and common cortical activations for multimodal semantic categories. *Cognitive, Affective, and Behavioral Neuroscience, 6*(3), 214-222.

Goldberg, R., Perfetti, C., Fiez, J., & Schneider, W. (2007). Selective retrieval of abstract semantic knowledge in left prefrontal cortex. *Journal of Neuroscience, 27*(14), 3790-3798.

Hinton, G. E. (1989). Connectionist Learning Procedures. *Artificial Intelligence, 40*, 185–234.

Lambon Ralph, M., Howard, D., Nightingale, G., & Ellis, A. (1998). Are living and non living category-specific deficits causally linked to impaired perceptual or associative knowledge? Evidence from a category-specific double dissociation. *Neurocase, 4*, 311–338.

Lambon Ralph, M., Patterson, K., Garrard, P., & Hodges, J. (2003). Semantic dementia with category specificity: A comparative case-series study. *Cognitive Neuropsychology, 20*(3), 307–326.

McClelland, J. (2013). *Explorations in parallel distributed processing: A handbook of models, programs, and exercises.*

McClelland, J., Rumelhart, D., & the PDP research group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Volume II.* MA: MIT Press.

McRae, K., Cree, G., Seidenberg, M., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Bahvior Research Methods, Instruments, & Computers, 37*(4), 547-559.

McRae, K., de Sa, V., & Seidenberg, M. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General, 126*, 99–130.

Merck, C., Jonin, P., Vichard, H., Boursiquot, S., Leblay, V., & Belliard, S. (2013). Relative category-specific preservation in semantic dementia? Evidence from 35 cases. *Brain & Language, 124*, 257–267.

Patterson, K., Nestor, P., & Timothy, T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Memory Systems: Reviews, 8*, 976-988.

Pinker, S., Mehler, J., Fodor, J., Pylyshyn, Z., Prince, A., Lachter, J., & Bever, T. (1988). *Connections and symbols.* MA: MIT Press.

Rogers, T., & McClelland, J. (2004). *Semantic cognitions: A Parallel Distributed Processing appproach.* MA: MIT Press.

Rosch, E. (1973). Natural categories. *Cognitive Psychology, 4*, 328-350.

Rosch, E. (1978). Principals of categorization. In E. Rosch, & B. Lloyd, *Cognition and Categorization* (pp. 27-48). NJ: Lawrence Erlbaum Associates.

Rossion, B., & Pourtois, G. (2004). Revisiting snodgrass and vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception, 33*(2), 217-236.

Rumelhart, D., Hinton, G., & Williams, R. (1986). Learning representations by back-propagating errors. *Nature, 323*(6088), 533-536.

Rumelhart, D., McClelland, J., & the PDP research group. (1986b). *Parallel distributed processing: Explorations in the microstructure of cognition. Volume I.* MA: MIT Press.

Samson, D., & Pillon, A. (2003). A case of impaired knowledge for fruit and vegetables. *Cognitive Neuropsychology, 20*(3-6), 373-400.

Sartori, G., Gnoato, F., Mariani, I., Prioni, S., & Lombardi, L. (2007). Semantic relevance, domain specificity and the sensory/functional theory of category-specificity. *Neuropsychologia, 45*(5), 966-976.

Seifritz, E., Esposito, F., Hennel, F., Mustovic, H., Neuhoff, J., Bilecen, D., . . . Di Salle, F. (2002). Spatiotemporal pattern of neural processing in the human auditory cortex. *Science, 297*, 1706-1708.

Siri, S., Kensinger, E., Cappa, S., Hood, K., & Corkin, S. (2003). Questioning the living/nonliving dichotomy: Evidence from a patient with an unusual semantic dissociation. *Neuropsychology, 17*(4), 630-645.

Smith, E., Shoben, E., & Rips, L. (1974). Structure and process in semantic memory: A feature model for semantic decisons. *Psychological Review, 81*(3), 214-241.

Szekely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., . . . Bates, E. (2005). Timed action and object naming. *Cortex, 41*(1), 7-26.

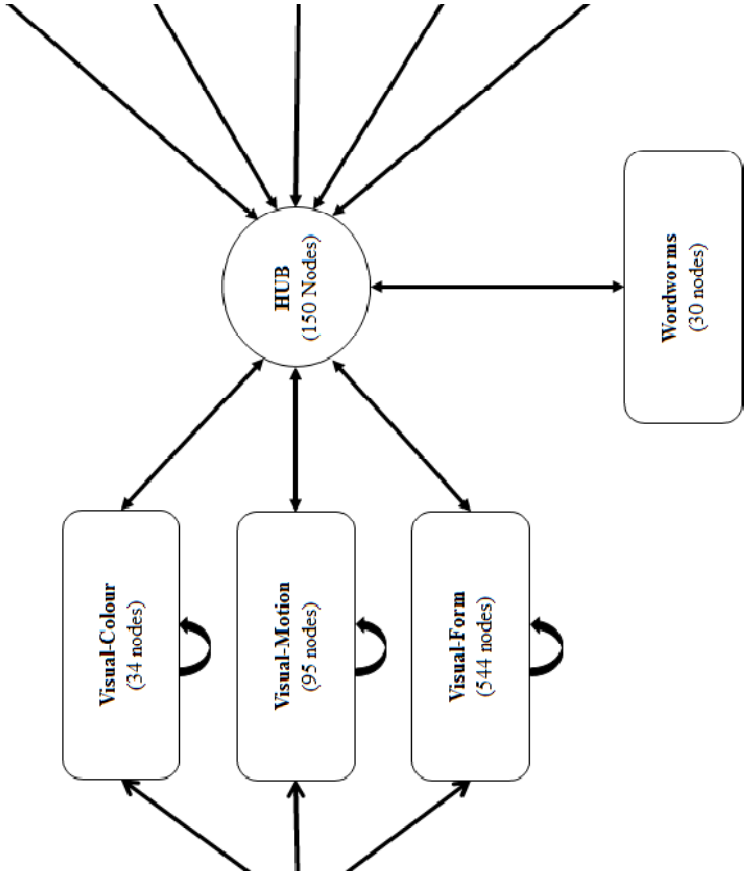
Tulving, E. (1972). Episodic and semantic memory. In E. Tulving, & W. Donaldson, *Organization of Memory* (pp. 381-402). NY: Academic Press.

Warrington, E., & Shallice, T. (1984). Category specific semantic impairments. *Brain, 107*, 829–854.

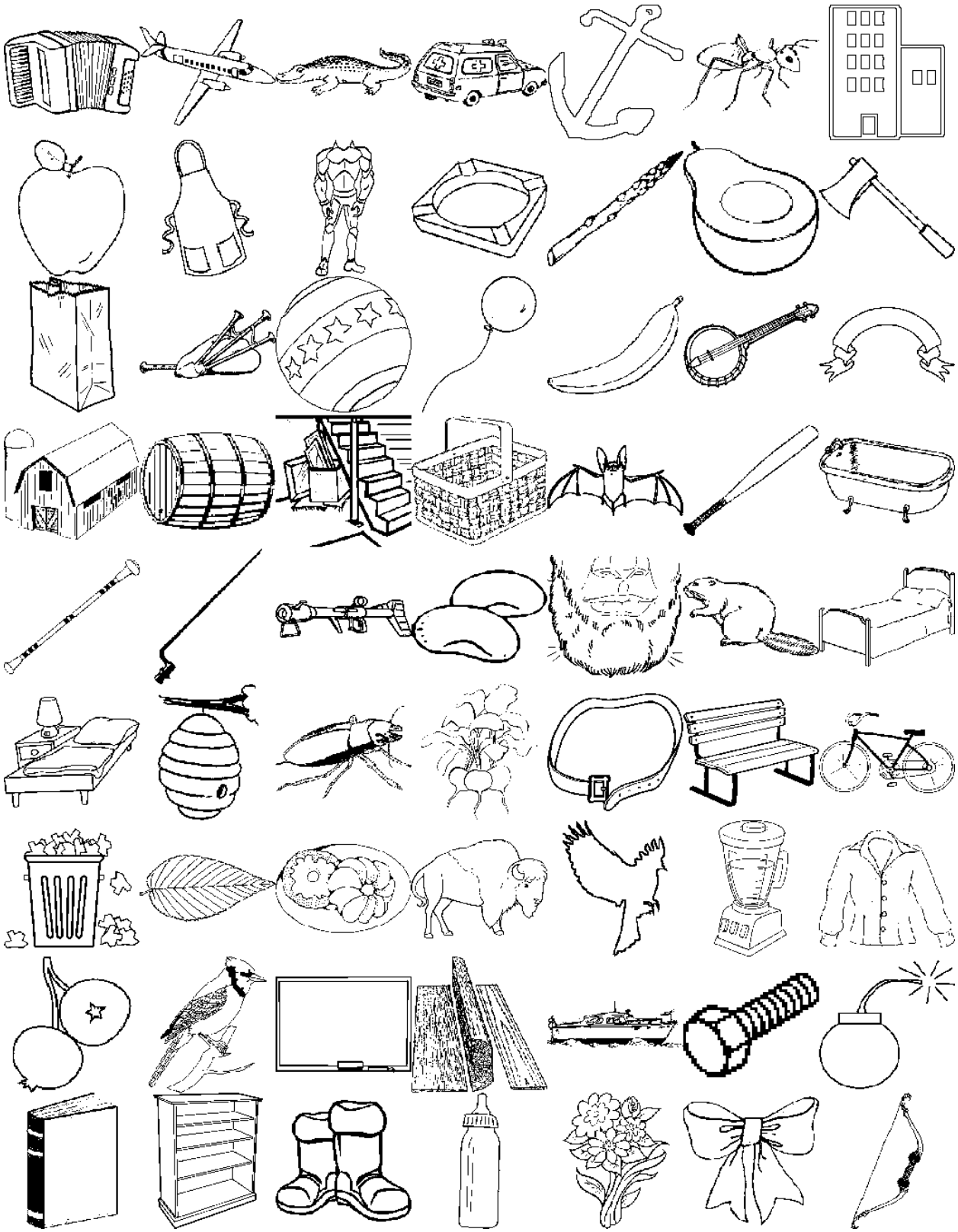
Wu, L., & Barsalou, L. (2009). Perceptual simulation in conceptual combination: Evidence from property generation. *Acta Psychologica, 132*, 173-189.

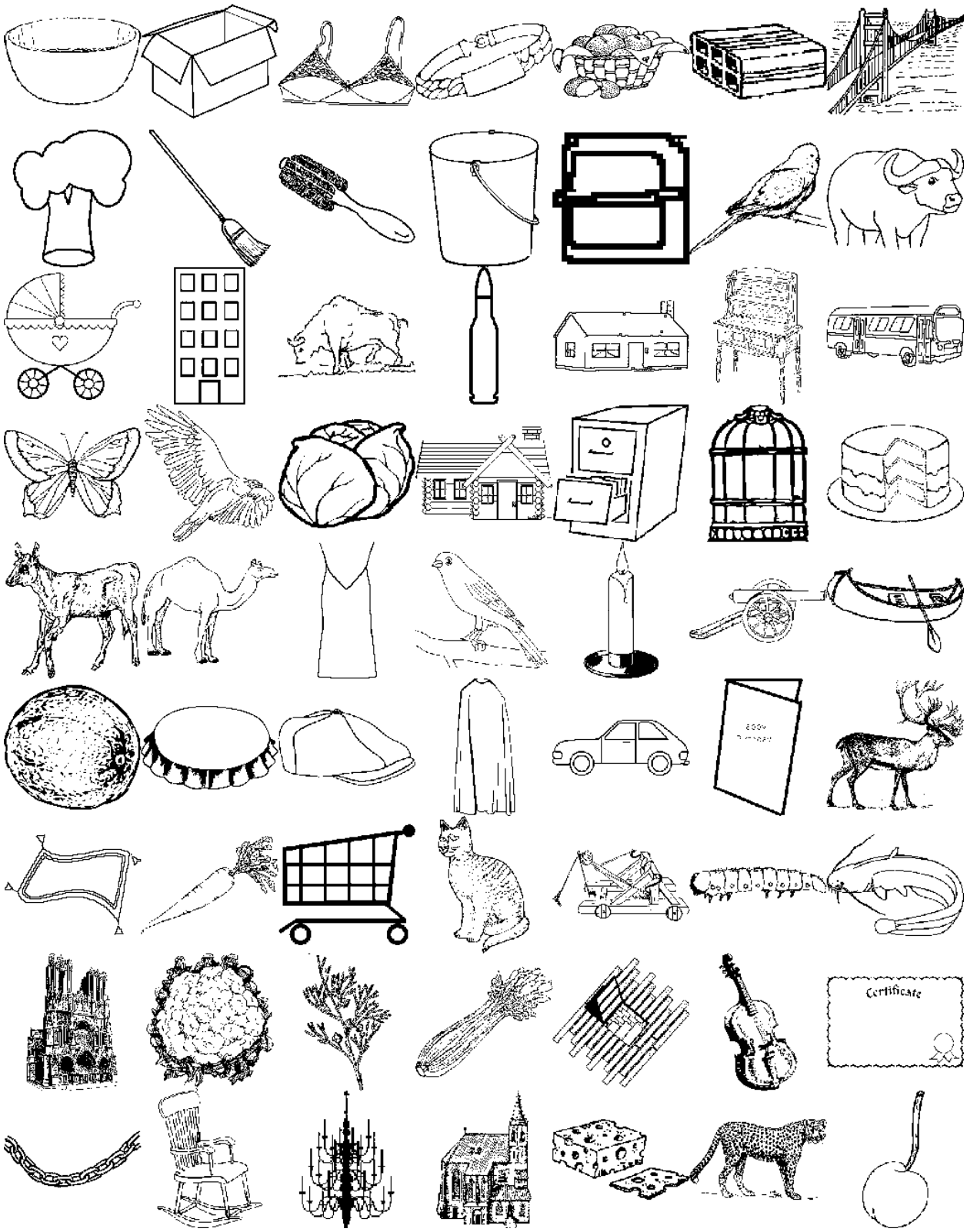
Zannino, G., Perri, R., Pasqualetti, P., Di Paola, M., Caltagirone, C., & Carlesimo, G. (2006). The role of semantic distance in category-specific impairments for living things: Evidence from a case of semantic dementia. *Neuropsychologia, 44*(7), 1017–1028.

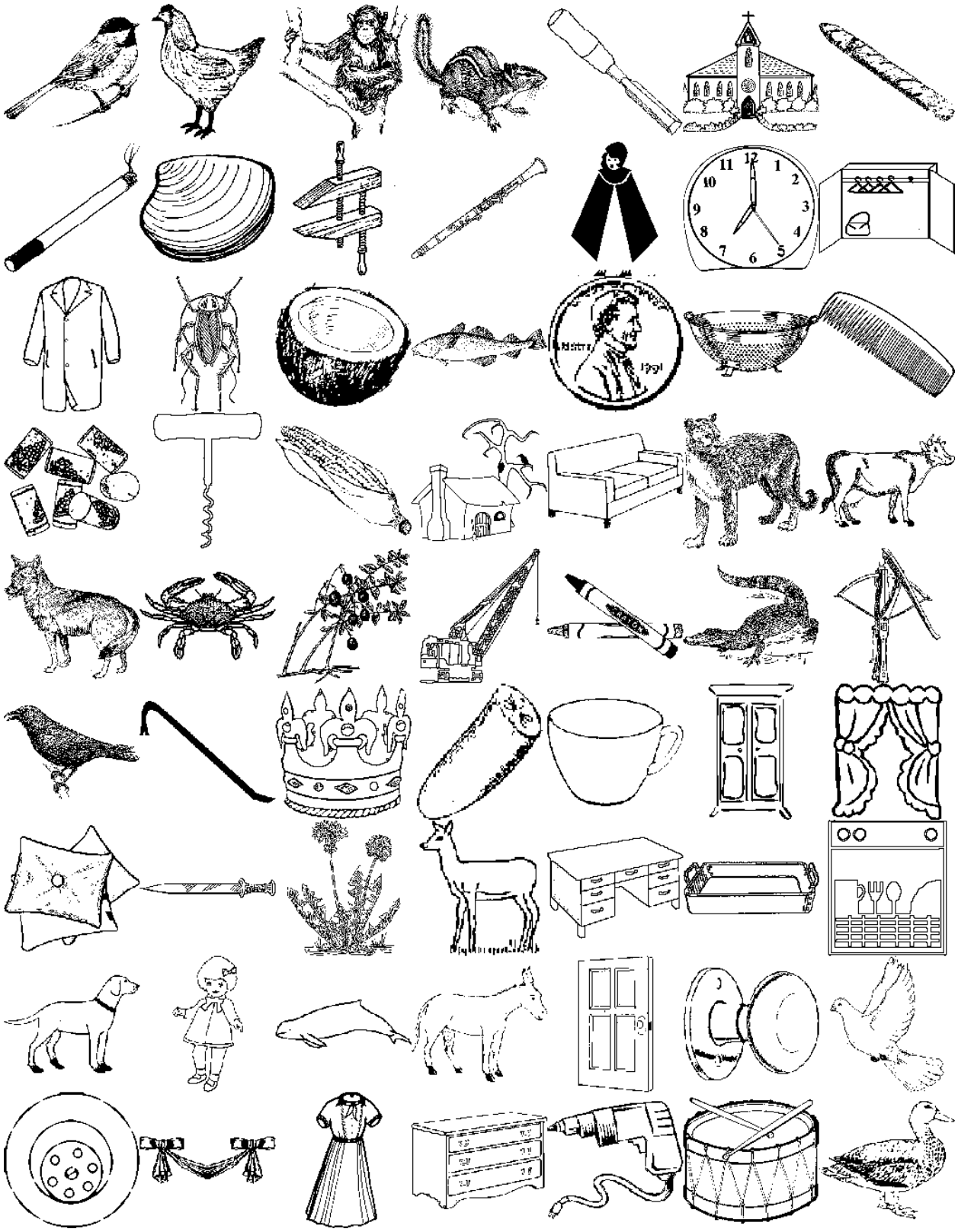
Appendix A: Non-Stylized Model Architecture

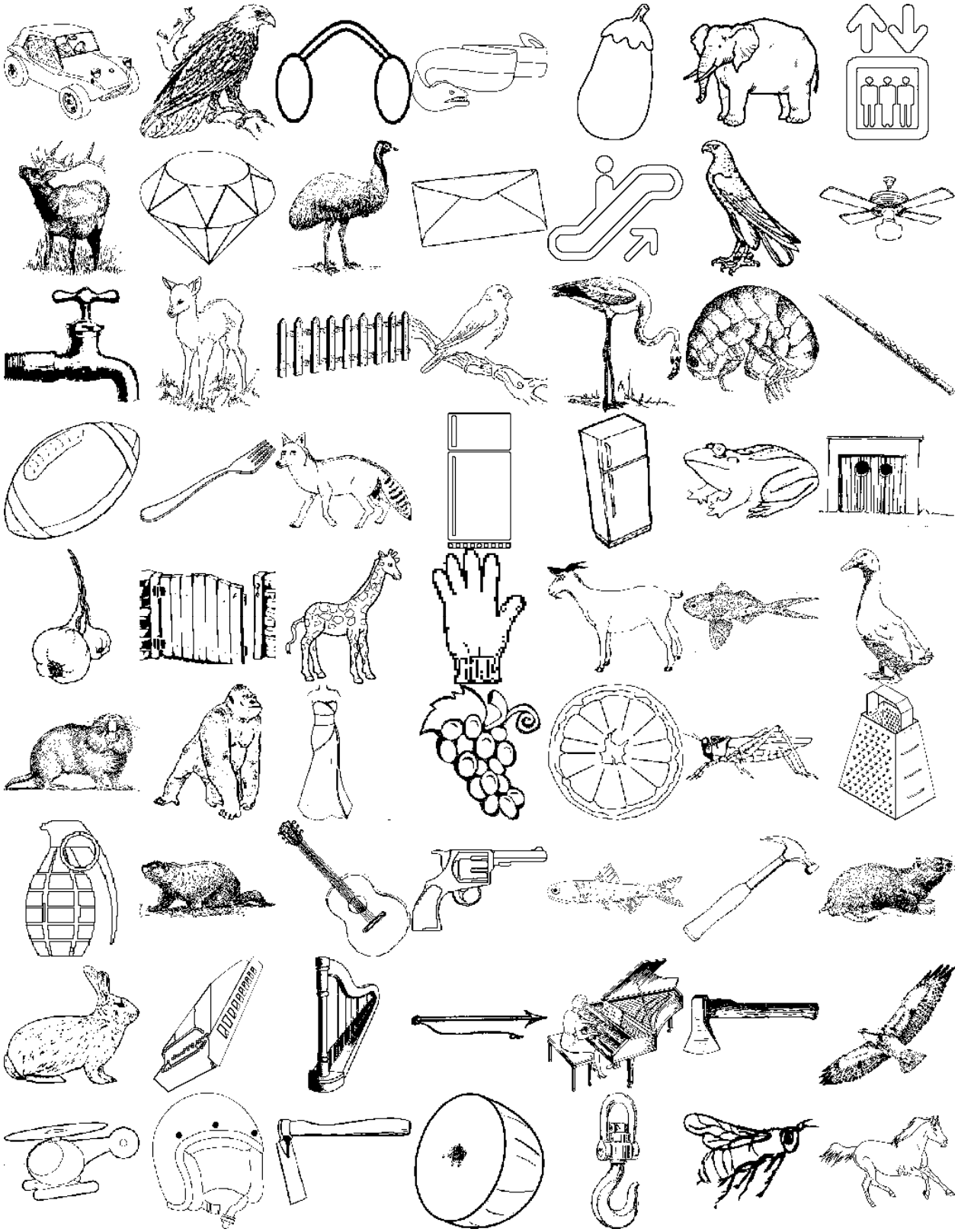


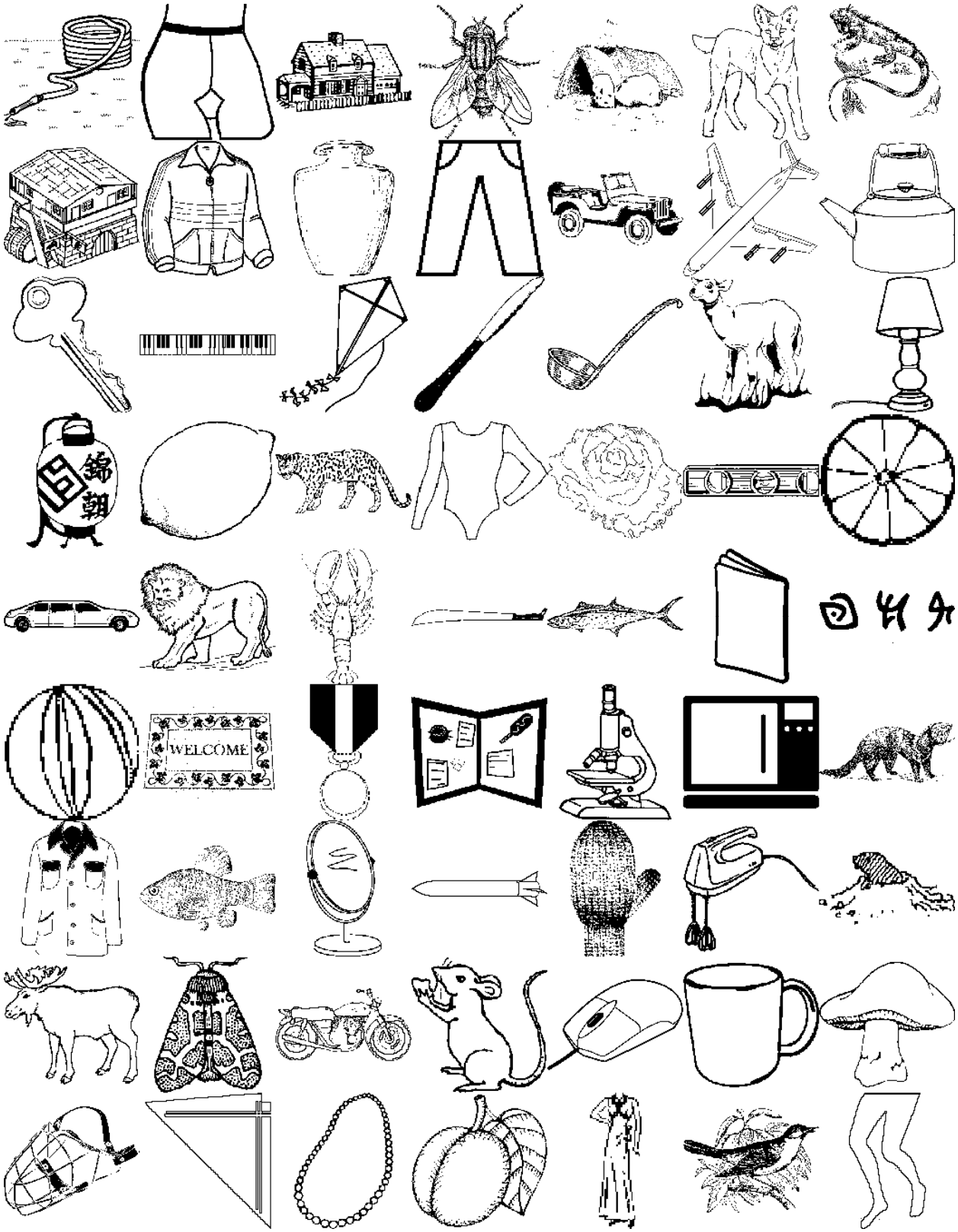
Appendix B: Line Drawings Prepared For Model Input




Certificate






WELCOME